Historic, Archive Document

Do not assume content reflects current scientific knowledge, policies, or practices.

a331
R44A7
c4

United States
Department of
Agriculture

**Agricultural
Research
Service**

ARS-81

June 1990

Proceedings of the International Symposium on Water Quality Modeling of Agricultural Non-Point Sources, Part 2

June 19-23, 1988
Utah State University
Logan, Utah

De Coursey, Donn G., editor. 1990. Proceedings of the International Symposium on Water Quality Modeling of Agricultural Non-Point Sources, Part 2. U. S. Department of Agriculture, Agricultural Research Service, ARS-81, 423-875 pp.

The papers, including references and tables, presented here were provided on magnetic media by the authors. They were reformatted for uniform spacing and heading style and then reprinted using a common print font. A few of the papers were retyped because the authors were not able to provide a magnetic media version. Figures are reproduced essentially as provided by the author. Queries regarding specific papers should be referred to the authors. The views expressed by the conference participants are their own and do not necessarily represent the views of the U.S. Department of Agriculture. Mention of trade names does not imply recommendation or endorsement by the Department.

The papers presented in this proceedings are in two groups: invited manuscripts on specific topics and poster presentations. The invited papers, which make up the bulk of the papers published here, were treated as technical journal papers and received peer and extensive editorial review. The contributed poster papers received editorial review.

United States
Department of
Agriculture

**Agricultural
Research
Service**

ARS-81

Proceedings of the International Symposium on Water Quality Modeling of Agricultural Non-Point Sources, Part 2

June 19-23, 1988
Utah State University
Logan, Utah

Donn G. DeCoursey, Editor

Sponsored by:
International Science and Education Council
Office of International Cooperation and Development, USDA
Agricultural Research Institute
Agricultural Research Service, USDA
Forest Service, USDA
National Agricultural Chemicals Association
National Association of State Universities and Land-Grant Colleges
Soil Conservation Service, USDA
Utah Agricultural Experiment Station
Utah Water Research Laboratory

CONTENTS

Page

Part 2 - User's Perspective

Surface Water Quality Models for Planning, Design and Operational Management
P.G. Whitehead, L. Somlyody and G. Van Straten....423

Problems with Surface Water Models from a User's Perspective
K. Thornton, C. Stalnaker and K. Baun....447

Discussion....457

Technical Session 5 - Integrated Models

Environmental Isotope Tracer Studies of Catchment Processes: Tools for Testing Integrated Water Quality Models
M.G. Sklash, I.D. Moore and G.J. Burch....459

Multimedia Modeling of Toxic Chemicals
Y. Onishi, L. Shuyler and Y. Cohen....479

Discussion....503

Technical Session 6

Part 1 - Spatial Variability and Scale

Effects of Spatial and Temporal Variability on Water Quality Model Development
D.A. Woolhiser, W.A. Jury and D.R. Nielsen....505

Effect of Spatial Variability Upon Subsurface Transport of Solutes from Nonpoint Sources
G. Dagan, D. Russo and E. Bresler....523

Discussion....549

Part 2 - Prediction/Comparison - Simple/Complex

Complexity and Uncertainty in Predictive Models
K.J. Beven and A.J. Jakeman....555

Selection, Application, and Validation of Environmental Models
A.S. Donigian, Jr., and P.S.C. Rao....577

Discussion....601

Part 3 - Risk Analysis/Confidence Limits

Analyzing Statistical Properties of Nonpoint-Source Water Quality Variables
J.D. Salas and J.C. Loftis.....605

Stochastic Aspects of Water Quality Modeling for Nonpoint Sources
E.J. Plate and L. Duckstein.....631

Discussion.....655

Technical Session 7

Part 1 - Management Improvements

Root Zone Processes and Water Quality: The Impact of Management
B.E. Clothier.....659

Management Improvements in Water Quality Models
D.G. DeCoursey and J.S. Schepers.....685

Discussion.....695

Part 2 - Manager's Perspective

A Management Perspective on Water Quality Models for Agricultural Nonpoint-Sources of Pollution
R.R. Shaw and J.W. Falco.....697

Water Quality Modeling: The Manager's Perspective
D.G. Jamieson, P.E. O'Connell and G. de Marsily (no paper available)

Discussion.....705

Technical Session 8

Part 1 - Software Engineering

The Potential Impact of Software Engineering on Water-Quality Modeling
D.G. Jamieson and K. Fedra.....707

Object-Based Software Development
S.N. Woodfield.....715

Discussion.....727

Part 2 - Parameter Identification

Application of Parameter Estimation Techniques to Solute Transport Studies
M.Th. van Genuchten, S.M. Gorelick and W.W-G. Yeh....731

Parameter Identification
U. Hornung....755

Discussion....765

Part 3 - Data Bases

Use of Soil Survey and Other Data Bases in the Modeling of Leaching from Agricultural Sources
A. Breeuwsma and J. Bouma....769

Data management for Water-Quality Modeling Development and Use
A.M. Lumb, R.F. Carsel and J.L. Kittle, Jr. 785

Discussion....799

Extended Abstracts of Poster Presentations

A Model for Simulation of Nitrogen Dynamics in Soil and Nitrate Leaching
L. Bergström, P.E. Jansson and H. Johnsson....803

The Soil-CREAMS Model to Simulate Soil and Chemical Losses from Agricultural Areas
P.F. Botterweg....809

RZWQM - A Model for Simulating the Movement of Water and Solutes in the Root Zone
D.G. DeCoursey and K.W. Rojas....813

Opus: An Advanced Simulation Model for Nonpoint-Source Pollution Transport at the Field Scale -- An Overview
V.A. Ferreira and R.E. Smith....823

A Simple Method for Assessing Pesticide Leachability
D.I. Gustafson....835

Mathematical Modeling of Sediment Loop Ratings: Problem and Literature Review
D.R. Jackson....841

Mathematical Model of Sediment Loop Ratings: Hydraulic Model
D.R. Jackson....847

Mathematical Modeling of Sediment Loop Ratings: Triangular and Gamma Function Model Development
D.R. Jackson....853

Mathematical Modeling of Sediment Loop Ratings: Triangular and Gamma Function Model Results
D.R. Jackson....859

A Model of Water and Solute Movement in Structured Soils
N.J. Jarvis....865

Integrated Assessment of Erosion and Nonpoint-Source Pollution in a Palouse Watershed
T. Prato and M. Brusven....871

Prediction of Chemical Transport in Agricultural Runoff: An Integrated Process Approach
A.N. Sharpley, S.J. Smith, L.R. Ahuja and G.C. Heathman....875

SURFACE WATER QUALITY MODELS FOR PLANNING, DESIGN AND OPERATIONAL MANAGEMENT

P. G. Whitehead[1], L. Somlyody[2] and G. Van Straten[3]

ABSTRACT

A review of model use in planning, design and operational management is presented for the U.K. and Continental Europe. Several applications are presented of river basin management and water body modeling and control. Common modeling problems are investigated and the key areas for future research identified.

INTRODUCTION

The purpose of this paper is to provide a state-of-the-art review of the use of surface water quality models in planning, design and operational management, particularly in relation to the control of agricultural nonpoint- source pollution. Before considering the models available it is important to define what is meant by point- and nonpoint-source pollution.

Point sources are discharges whose total flow is conveyed in a well-defined channel. Typical examples are municipal and industrial discharges of waste- water, storm water culverts and sewer overflows. These all possess the property that the total load of pollutants can be determined by sampling and flow measurement at the point of entry to a receiving watercourse.

Diffuse or nonpoint-sources are laterally extended discharges where the total flow cannot be measured or sampled directly or even readily observed at a single point. The processes which pollute these discharges may be discrete or dispersively spread over the catchment. Thus the essential distinction between point and diffuse discharges is the ability to remove the pollutant from the former. This contrasts with the methods of control available for diffuse sources which can only be regulated by curtailing an activity, which if uncontrolled, would give rise to pollution (e.g., careless spraying of crops with pesticide, or an excessive use of fertilizers on agricultural land).

POLLUTANTS FROM AGRICULTURE

Over the past 30 years there has been an intensification of the production of many crops and farm animals. This has been brought about by: increased use of agrochemicals (fertilizers and pesticides); increased mechanization and increased specialization. Most pollutants arise from the increased use of agrochemicals.

[1]P.G. Whitehead, Head of Environmental Hydrology, Institute of Hydrology, Wallingford, U.K.

[2]L. Somylody, Director, Research Centre for Water Resources Development, VITUKI, Budapest, Hungary.

[3]G Van Straten, Assistant Professor, Department of Chemical Technology, University of Twente, Enschede, The Netherlands.

Fertilizers

The application of chemical fertilizer and manure has increased substantially in most developed countries over the last 50 years. The effect of increased nitrogen usage is reflected in increasing concentrations of nitrates in rivers, lakes and groundwaters. The annual mean concentration of NO_3-N in the River Thames at Walton in 1950 was around 5mg/1, but in 1976 the mean concentration was around 15mg/1. A doubling in the nitrate concentration from 5 to 10mg/1 has been observed in borehole water from the Triassic sandstone of Worcestershire (UK). Part of the explanation for these increases is the extra quantities of nitrogenous fertilizers used by farmers, but the change from grassland to arable land is also implicated. There is adequate evidence to show that the drainage water from under arable land contains a great deal more nitrate than that from grassland. The relevance of this to pollution is the recommendation of World Health Organization not to supply water in excess of 11.3 mg No_3-N/liter because of the possibility of causing methemoglobinemia in infants. A further factor is the role of inorganic nitrates in stimulating excessive plant growths (eutrophication) in lakes and reservoirs. However, this can only occur if inorganic phosphates are also present.

In most undeveloped upland catchments, the concentration of inorganic phosphate in waters draining them is very low, usually a few micrograms per liter. Inorganic phosphate is retained by most agricultural soils and evidence points to very little fertilizer-derived phosphorus moving into drainage waters or deep aquifers in the way that nitrate does, because of its tendency to absorb to the soil. The amount of phosphorous in drainage water is controlled by soil type rather than rates of application. It is only on soils containing little or no clay component, e.g. light sandy or peat soils, that substantial losses of phosphorus occur. It is significant to note than increased afforestation in upland areas with mainly peat soils may well bring about significant increases in phosphorus concentration in waters draining from such areas. This is particularly important because the upland regions are also the main gathering grounds for water, and eutrophication problems in upland reservoirs are likely to occur. Concentrations as low as 0.01 mgP/liter have been shown to stimulate excessive growths of algae in lakes and reservoirs.

Pesticides

The use of chemicals to control weeds, pests and diseases has increased enormously since 1950. The term pesticide embraces insecticides, acaricides, nematicides, molluscicides, herbicides, fungicides, soil fumigants and growth regulators. So far as water quality is concerned consideration has to be given to the application and movement of pesticides, their degradation products or their persistence in the environment and the importance of these to ecology, including humans.

Probably the most persistent pesticides are the organochlorine compounds whose traces are to be found throughout the world. During their passage through the food chains these chemicals are concentrated in the fatty tissues and for this reason, their use is banned in many countries. The rate of introduction of new pesticides is slowing down, largely because of the great cost and increased commercial risk. Manufacturers now look for pesticides which will control a broad spectrum of organisms rather than highly selective ones for particular pests. Such developments will inevitably lead to greater problems in freshwaters when accidental contamination occurs. From time to time dramatic episodes are caused by the unintentional entry of pesticides into watercourses. Often these are caused by ignorance, carelessness or mechanical failure and these incidents should not be regarded as indications of the real dangers from the use of such chemicals.

Intensive Livestock Farming

Increasing the intensity of livestock farming increases the problem of animal slurries as potential pollutants. Lowland beef and dairy cattle are normally housed for part of the year, and usually the

slurry that accumulates during that period can be disposed of to the grazing land. Pigs and poultry may be housed in intensive units and may pose serious pollution problems because the conditions can demand more land than is available for slurry disposal. Animal slurry is characterized by its high Biochemical Oxygen Demand (BOD), high concentrations of nitrogenous compounds, phosphates and suspended solids. Even when successfully applied to land, the slurry may create problems because disposal is often necessary during the wet season when soil leaching and wash- off is most intensive.

Associated with the more intensive livestock farming is the increase in the use of silage as a winter feed, and silage liquors with a BOD in excess of 10^4 mg/liter occasionally cause local pollution problems.

From a management perspective there are clearly major sources of nonpoint pollution and it is important to bear these in mind when addressing the use of models to control pollution.

REVIEW OF MODEL USE

Figure 1 depicts major differences between point- and nonpoint-source pollution problems and their modeling presented incorporating major factors and processes affecting water quality, control possibilities and various elements and stages of management and policy making.

The figure indicates the place and role of various descriptive, planning and management models for use in watershed studies and water bodies such as lakes and reservoirs.

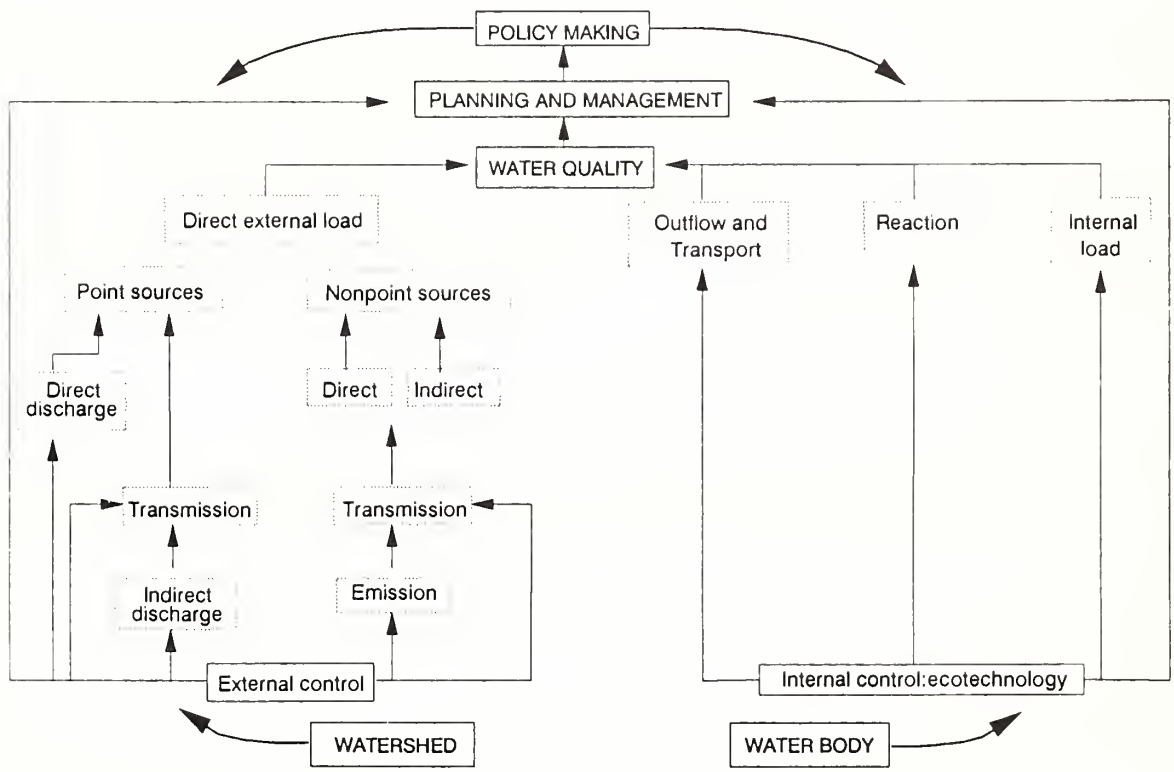


Figure 1.
Major elements of water quality management.

The review of model use has been undertaken from three different perspectives: the U.K., Eastern Europe and Western Europe. The division is obviously somewhat artificial. However, the three reviews are presented separately to indicate the differing approaches and the range of models and techniques available.

The U.K.

Some 10% of the length of U.K. rivers (a total length of 3,790 km in England and Wales) are deemed to be of unsatisfactory quality. These rivers only support sporadic fish populations and are generally associated with areas of urban development. However, in recent years there has been increasing levels of pollution from nonpoint pollution derived from catchment land-use changes and agriculture. Table 1 shows a summary of water quality problems identified by the U.K. Water Industry. Nonpoint-source pollution is identified by the majority of the water authorities as being of considerable importance. However, their modeling efforts have been largely confined to the effects of pointsource pollution such as effluent discharges on river systems. These models such as QUAL II, TOMCAT and SIMCAT (see Crabtree et al. 1986) have been used to establish effluent consent conditions. These planning/design models provide steady-state estimates of instream water quality and in the case of TOMCAT and SIMCAT have been enhanced using statistical techniques to provide 95-percentile levels as required by the UK water pollution regulations. Models for operational management and control have been developed by a number of groups in the U.K. including Whitehead et al. (1979, 1981) and Beck et al. (1987). In the case of models developed by Whitehead et al. these have been linked via telemetry to on-line river monitoring schemes and employed to forecast in real time the impact of pollution along the river (see Whitehead et al. 1984). Such models can be used to investigate the relationship between point and nonpoint pollution as indicated in figure 2. This shows the simulation of nitrogen along the Bedford Ouse

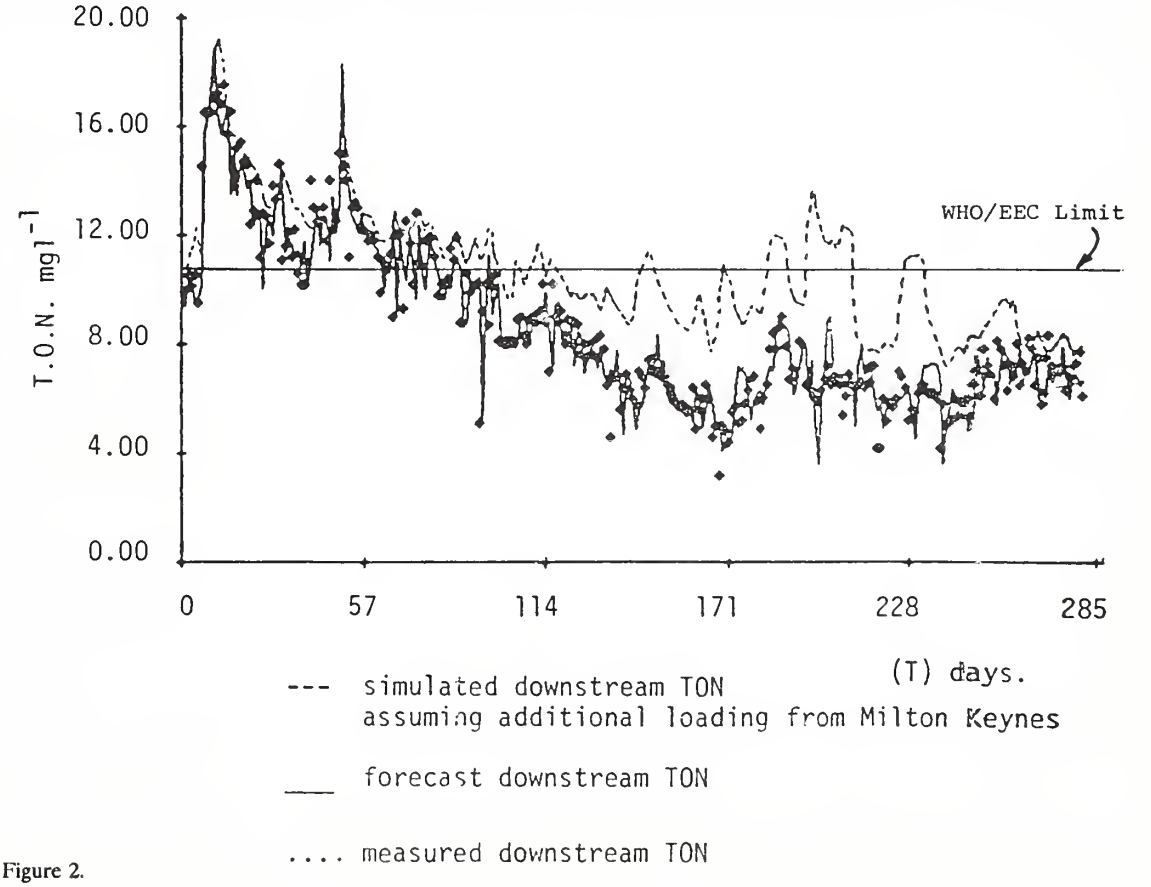


Figure 2.
TON (Total Oxidized Nitrogen) Variations in the Bedford Ouse in 1974.

Table 1.
Rank (by perceived importance) of river water quality problems identified by the U.K. Water Industry.

	Continuous Discharges	Non-Point Pollution	Urban Runoff Storm Over-Flows	Pollution Incidents
REGIONAL WATER AUTHORITIES IN ENGLAND AND WALES				
Anglian	1	2		
Northumbrian			2	2
North West	2	3	1	
Severn Trent	4	3	1	
Southern		1	2	
South West	2	1		
Thames	1	3	2	
Welsh		3	2	1
Wessex	2	1	3	
Yorkshire	1	3	2	
RIVER PURIFICATION BOARDS IN SCOTLAND				
Clyde			2	3
Forth	1	3		
North East		2		1
Tay	3	2		

and the impact of both rural runoff and a major effluent discharge on the river. In terms of nitrogen the point-source effluent has minimal impact during high flow winter conditions because nonpoint sources of nitrogen dominate river quality. It is only in low flow summer conditions that the point source pollution becomes significant. These nonpoint sources of nitrogen from agricultural activities has been addressed in a more detailed study undertaken by the Thames Water Authority and this is described in detail below.

With regard to pesticides there has been little modeling research in the U.K. and only recently have modeling applications been considered by Water Authorities. A research project established by the Water Research Centre in collaboration with Anglian Water aims to evaluate pesticide levels in groundwater. The Institute of Hydrology has also established pesticide projects with Wessex Water and Welsh Water to evaluate pesticide movement during storm events. In the latter studies hydrochemical models are being developed to evaluate flow paths and processes; pesticide models such as TOXIWASP, EXAMS AND CREAMS are being evaluated.

With regard to eutrophication models these have been developed largely for reservoir systems (see Steel 1975) and there has been almost no applications of the EPA type reservoir models by U.K. Water Authorities. The approach of Steel has been to estimate algal growth given nutrient conditions and reservoir characteristics, and to use the models to design reservoir operating rules.

Very few algae models have been developed for rivers,although Whitehead and Hornberger (1984) developed a model of algal growth and transport for the River Thames system.

In general, model use by U.K. water authorities has been driven by practical problems requiring a solution. Thus principal applications have been for:

a) setting effluent consent conditions;
b) forecasting flow and quality along critical reaches of river; and
c) designing river and reservoir operating rules.

It is only very recently that nonpoint sources of pollution have become a dominant factor and there has been a demand for models.

Eastern Europe

A major problem in Eastern Europe is the increasing levels of nitrate in rivers due to the increased loads of point-source effluent and nonpoint agricultural sources. As an example, the nitrate concentrations in the Danube, downstream of Budapest, has increased from 2-3 mg/l to 18-20 mg/l over the past 25 years. QUAL II has been employed to model nutrients and oxygen with the primary use for evaluating various planning and management tasks. The most recent standardized (and well documented) version of QUAL II can be run on an IBM-compatible PC and in principle allows comparisons from all over the world because of the unified character of the model. This version of QUAL IIe has been extensively used (Brown and Barnwell 1986).

Previous versions of the model were implemented for small- and medium-size rivers, but no application is known for larger rivers such as the Rhine or Danube, for which the assumption of complete mixing in cross-sections may not be applicable (Somlyody 1978b).

Eutrophication cause by pollutants has been identified by several researchers and river managers as the most serious problem of lakes and reservoirs. This is why the literature of modeling is richest in this field (see e.g. Scavia and Robertson 1979, Straskraba and Gnauck 1985).

Ecological models may include different fractions of phosphorus and nitrogen as state variables, and the biomass of dominant species (phytoplankton, zooplankton etc). The number of state variables ranges between 3 and 40, while that of model parameters is 10-100's (see e.g. Straskraba and Gnauck 1985, Somlyody and van Straten 1986), depending on complexity, modelling philosophy and data availability. A majority of models are deterministic simulation models assuming a fixed structure of the ecosystem. With respect to vertical segmentation, three layers, the epilimnion, hypolimnion and sediment can be distinguished. Horizontal segmentation primarily depends on the load pattern and hydrodynamics. Further classification of models can be made on the basis of degree and extent of applying techniques such as sensitivity and uncertainty analysis, parameter estimation, and model structure identification.

Several examples can be mentioned for ecological models of lakes and reservoirs. The model family AQUAMOD was developed and applied for different water bodies in GDR and CSSR (Straskraba and Gnauck 1985).

The most complex version, AQUAMOD 3, incorporates three layers and eight state variables (phytoplankton and phosphate-phosphorus in epi- and hypolimnion, respectively, filtrating zooplankton in epilmnion, organic matter, dissolved phosphorus and adsorbed phosphorus in sediment).

Another (two-layer) dynamic ecological model, called SALMO, the applicability of which was tested for quite different lakes and reservoirs, was developed at the Hydrobiological Laboratory of

the Water Resources Department of the Dresden University of Technology (Benndorf and Recknagel 1982). The model incorporates four state variables for both layers (dissolved orthophosphate, two algal groups and zooplankton). Despite the low number of state variables, the model involves about 80 parameters, which are determined on the basis of laboratory and in situ experiments, and literature. A similar model, more complex toward the role of the sediment, was developed by Kozerski and Dvora'kova' (1982). Further examples can be found, e.g., in Vavilin et al. (1979) and Svirezhev et al. (1979).

The behavior of shallow lakes is more complex than deep lakes because of the relatively thick photic zone, the lack of hypolimnion, the influence of wind- induced circulation and sediment resuspension and the dominant role of external (cross-correlated) meteorologic and hydrologic factors (such as solar radiation, temperature, wind, precipitation, and various load components).

Lake Balaton, one of the largest shallow lakes of the globe subject to artificial eutrophication was studied intensively, with one of the objectives being to work out and compare various ecological models. Three model families of 4-10 state variables and 30-50 parameters were developed (Somlyody and van Straten 1986, Leonov 1985). Two of them were phosphorus cycle models, while the third incorporated nitrogen and blue-green algae as one of the four algal compartments. Zooplankton was excluded because of its low biomass and negligible role on nutrient cycling. Hydrodynamics and mass exchange were accounted for on various levels of complexity, starting from a horizontally 2- D circulation model, the output of which was aggregated into 1-D longitudinal dispersion equations set for all state variables (the elongated shape of the lake supports the rationality of this approach) or into further simplified equations for four interconnected, completely mixed reactors (lake basins). Wind-induced sediment resuspension and its influence on light conditions were also studied and modeled (Somlyody 1982 and 1986a).

Parameters of ecological models outlined above can be partly determined from experiments; however, calibration can never be avoided. Because of complexity, nonlinearities and data scarcity, this is done for most of the cases by "fine-tuning", and not by formal parameter estimation techniques. Little progress has been achieved in this field during the past ten years and perhaps the hypothesis-testing procedure involving also the estimation of parameters (Hornberger and Spear 1980) seems to be the only promising method in this respect (see Somlyody and van Straten 1986 for an application to Lake Balaton).

The state-of-the-art of ecological models is perhaps best characterized by the general observation that model calibration is often successful. However validation in a stricter sense frequently leads to failures.

Ecological models are primarily developed by scientists whose main objective is understanding. Accordingly, the definition of practical objectives is often missing, and the usage of existing models for planning and management is sometimes beyond the existing possibilities. One of the exceptions is the Lake Balaton case, for which a stochastic, planning model was derived and incorporated with a watershed model into an optimization model. Another exception can be observed in the field of ecotechnology in which models such as AQUAMOD or SALMO are put into a dynamic optimization framework (Straskraba 1985, Recknagel et al. 1986). In such a manner not only internal control alternatives, but the "best" combination of external and internal measures can be analyzed (see fig. 1).

Within Eastern Europe there has been development of watershed models to assist in the management of nonpoint-source pollution.

At an early stage in nutrient model development it was noted that for nonpoint- source pollution:

(1) not only the water pathways should be described, but also nutrient pathways;
(2) it is not yet known how to transit from the often "unrealistically" small field size to larger areas; and
(3) the data requirement is extremely large.

These problems with structural models led to the application of empirical models (primarily in relation to phosphorus). Most of them compute the volume of surface runoff and the sediment yield for a precipitation event (e.g., on the basis of equations like the SCS curve number approach and the Universal Soil Loss Equation). The dissolved P loading of the event is obtained by multiplying the runoff volume by the dissolved P concentration of overland flow; the particulate P loading is the product of the sediment yield and the average adsorbed P concentration of the sediment. The annual total load is obtained by summing the loads of events over the year. The application of such an empirical model for the Tetves subwatershed of Lake Balaton (70 km^2) is given in Bogardi and Duckstein (1978).

In general, it is felt that the advantage of empirical models is that they include parameters which express the influence of cropping management. The determination of these parameters is, however, quite subjective and therefore the advantage is only illusory. Thus the conclusion is that the application first requires an overall practical knowledge of the watershed under study.

It is noted that the partially calibrated model of Bogardi and Duckstein (1978) was later incorporated in a multiobjective management model serving conclusions of methodological importance (Bogardi et al. 1983).

A perhaps even simpler but more complete and effective approach is based on unit areal loads and transmission coefficients. The procedure can be easily implemented in an expert system fashion on an interactive PC. Major steps are as follows (Somlyody and van Straten 1986):

(1) The "hydrologic tree" of the watershed is produced;
(2) imposed on the hydrologic tree, another segmentation is performed showing areas according to different types and amounts of point- and nonpoint-source pollution. On the basis of unit areal loads and river transmission coefficients (for which quite a lot of literature data exist and refinements are possible by field observations) a descriptive model can be obtained for the entire catchment (which should be calibrated by using measurements at the mouth section of the main river);
(3) the introduction of control alternatives (fertilizer use, erosion control, pre-reservoirs etc.) results in a planning model which distinguishes the origin and location of load components, as well as control measures;
(4) stochasticity and uncertainty can be incorporated in hydrologic, meteorologic and load determinations; and
(5) the planning model thus obtained can be incorporated into a management model.

The implementation of such a planning/management model was undertaken for the Tetves watershed divided into 17 sub-basins (Jolankai and Pinter 1986). Unit areal loads were considered as a function of runoff and 12 land use forms. The management model had the objective of minimizing the load subject to cost and technological constraints. The final result is a function between overall load removal efficiency and total cost, and the sequence of the 17 sub-basins in which the control actions should be performed to attain maximum cost effectiveness.

A different approach was developed for another pilot zone, the Rakaca watershed (~210 km^2) in Hungary. Here nitrogen pollution forms the major problem because the Rakaca reservoir serves as a basis for drinking water supply. A conceptual, farm-scale model (DISNIT) estimating the amount of those forms of nitrogen which can be displaced by water was determined. It considers in a simplified way the rooting zone's water and nitrogen balances, using empirical relationships.

The water components of the system include rain and irrigation water as inputs, and evapotranspiration, runoff and infiltration below the rooting zone as the outputs. For nitrogen, the system includes the nitrogen originating from natural processes in an aggregated form and nitrogen coming from agricultural practices as inputs, while vegetal nitrogen consumption and nitrogen carried away by water horizontally and vertically are the outputs. The empirical relationships constituting the model structure take into account the type of soil, the kind of crop and the effects of the various agricultural practices, but depend on average meteorological conditions.

The water model output consists of daily values of infiltration into and from rooting zones, as well as the momentary water content thereof, and the amount of evapotranspiration and runoff. The displaceable nitrogen (the amounts consumed by the nitrogen component output includes vegetation and removed by water from the surface and into the soil layers below the rooting zone) as well as the momentary displaceable nitrogen reserves.

Subsequently, the use of the farm-scale DISNIT model has been extended to the whole basin in order to compute the quantity of potentially displaceable nitrogen. Taking into consideration different cultivation practices, two types of soils, and five different crops planted, the amount of nitrogen loss to be expected has been calculated by Monte Carlo simulation. The management model developed (Pinter and Feher 1987) had the applicable cultivation practice of different plots as the decision variable, with one deterministic and two stochastic constraints to maintain the total quantity of displaceable nitrogen under a given level with minimal costs. On all plots the cultivation practice of contour tillage plus deep plowing proved to be preferable. Implementation of the proposed measures of pollution control and cultivation practice could decrease the annual total nitrogen load originating from surface runoff by approximately 20%. The development and further application of the above system of models is under way.

Western Europe

In Western Europe there is a general feeling among water authorities and water managers that computer modeling can be a valuable tool in solving management and planning problems. The role of modeling is, however, somewhat de-emphasized because for a number of problems relatively simple models can give sufficiently accurate answers. Also it should be noted that the European Community policy with respect to environmental legislation more or less by-passes the need for modeling in many instances. This is because legislation mostly concentrates on a system of levies and discharge permits, where the receiving water quality plays a much less important role than equal treatment of different polluters and fear for competition advantages. So, in assessing permissible waste loads there is no need for considering different alternatives, because the alternatives are restricted by law. On the other hand, in recent years public water management has been confronted with increasingly complex problems, for which straightforward solutions simply do not exist. This is, for example, the case in eutrophication management, where the feeling is that simple reduction of phosphorus point sources will not be sufficient to reclaim our freshwater lakes. Besides, there is a growing concern about pollution of the seas, especially the Mediterranean and the North Sea, which require proper assessment of the effects because remedial measures will be very costly.

AVAILABILITY OF COMPUTER MODELS

The manager (or researcher, for that matter) who recognizes the needs for a computer model now faces the difficult problem of knowing which potentially suitable model is available. In Europe, there is not a community agency that acts as a clearing house for environmental computer models. Neither are there such agencies in the individual nations, as far as we know.

Not knowing where to go, the manager might explore the scientific literature on mathematical modeling of water quality, in the hope that associated computer models are available or at least accessible. However, Western European research groups with publications in the open literature are quite limited. Table 2 summarizes the names of modelers in Western Europe quoted by Beck (1985) in his review on applications of mathematical models.

This does not, of course, mean that there are no more modelers or modeling activities, as will be substantiated in the next section. Only, there is little publicly available information on computer models derived from mathematical models, and even less on their application in practical management problems.

A Selection: Review of Water Quality Models in The Netherlands

The lack of freely available publications on water quality computer models makes it almost impossible to give a comprehensive overview of what is available. However, to demonstrate the extent of "subsurface" models we describe the situation in the Netherlands. This country is a favorable example, because the information is easy to access thanks to a series of reports on tools available for Environmental Impact Assessment prepared for the Dutch Ministry of Housing, Regional Planning and Environmental Management. The volume on surface water reviews the methods available and used by research institutions, consultancy firms and universities.

Table 2.
Western European research groups on water quality modeling cited by Beck (1985), and their countries of origin.

Origin	Modeling Group
Finland	Kinnunen et al.
Norway	Christophersen et al.
Austria (IIASA)	Fedra
Denmark	Jørgensen et al.
FRG	Hahn et al.
	Boes (1978)
	Stehfest
Italy	Rinaldi et al.
	Marsili-Libelli
Switzerland	Imboden et al.
Netherlands	Van Straten et al.
UK	Whitehead et al., Beck

For the present purpose the information is structured along the lines of figure 3, which gives a broad outline of the different subsystems in water quality modeling. Each of the subsystems requires inputs, which can either consist of available measurement data, or should be generated by running a model for one or more of the other subsystems; of course, the water system as a whole requires external inputs such as solar radiation, precipitation, wind, waste loads, and storm water. Here again, the choice is to use measured data, or to use generated data, based on some statistics or model (for instance from a watershed rainfall runoff quality model).

Computer models for the subsystems (noted "A"-"D" in fig. 3) on water movement and heat distribution are briefly summarized in table 3. A more detailed account is given for water quality components (numbered 1-6 in fig. 3) in Table 4. Models which inherently contained information specific to a particular geographic region have been omitted, since adaptation to other locations is neither recommended nor often possible. Some applications of these models are given in the next section of this paper.

SPECIFIC CASE STUDIES

The Thames Nitrate Model

In the Thames River Basin there has been a major change in landuse over the past forty years, as indicated in figure 4. Onstad and Blake (1980) estimated that the major land- use change from permanent grass to cereal crops together with significant use of nitrogen fertilizers released large loads of nitrate. These nitrate loads have been reflected in higher river nitrate concentrations, as shown in Figure 5.

Table 3.
Summary of computer models for water movement and heat distribution in The Netherlands.

SUB-SYSTEM	MODELED PROCESSES	AVAILABLE MODELS
A	2-D or 3-D currents in lakes and seas	FRIMO, ESTFLO, ODYSSEE, SIMONS
B	flows in networks of watercourses	NETFLOW, RIBASI, SOPWA, HYDRA, IMPLIC, DIWA
C	Thermal stratification in lakes	RESQUA, DYRESM, MITEMP, HEATEX STRATIF, BSTRAT, STRESS, STRABE
D (thermal)	{Near-field computations {of heat discharge effects {	MIT, PDS, STRAAL-3D
	{Far-field computations of {heat discharge effects	DFTHOR, TEMVER, KOWA

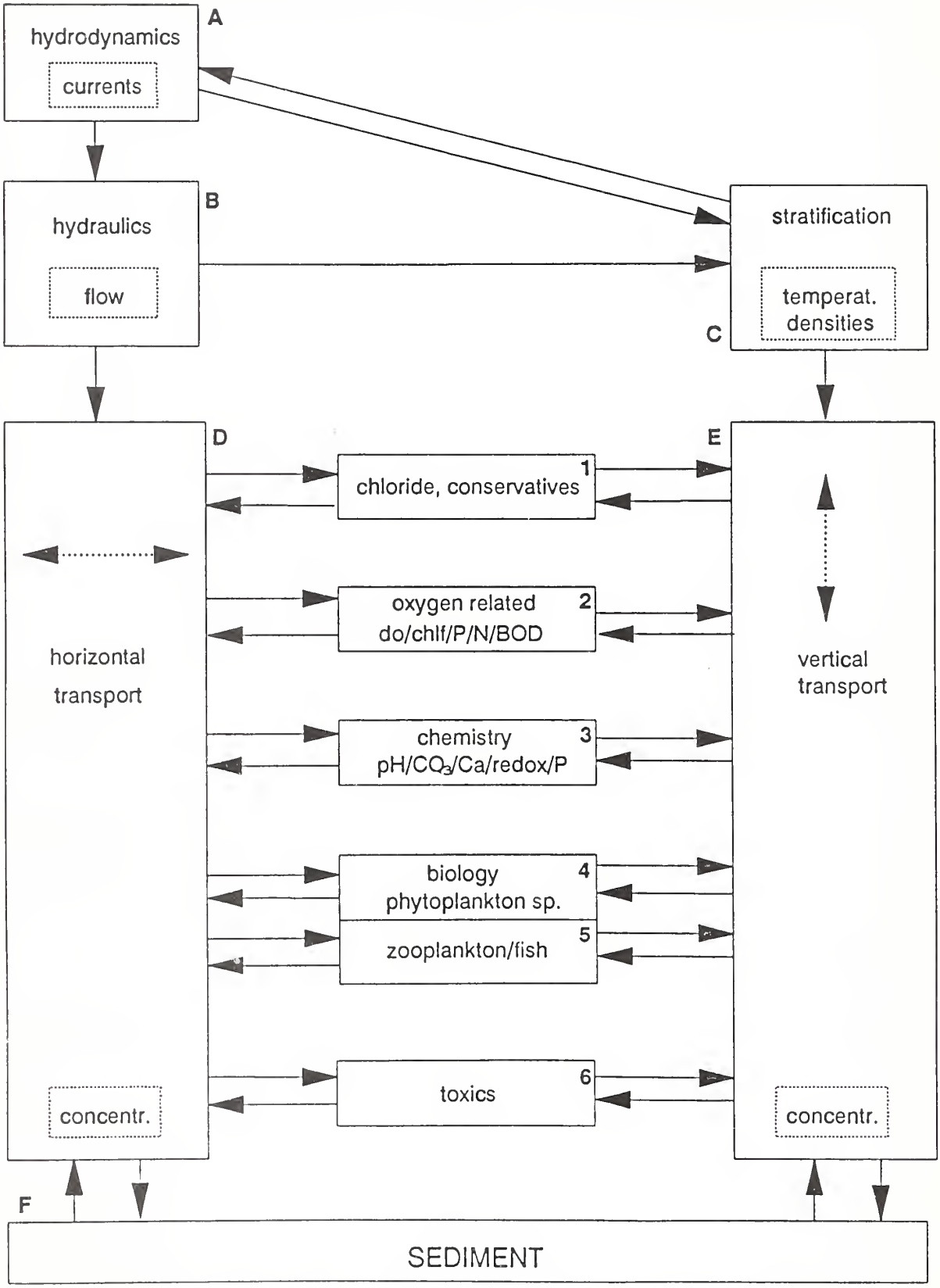


Figure 3.
Factors affecting water quality.

Table 4.
Survey of water quality models in The Netherlands.

MODEL	CASES[1]	INPUT[1]	REMARKS	WHO[2]
Node-link and network/stream models				
ARIADNE	D1	B	steady state	DHL
ABOPOL	BD1		instationary distribution on steady flow	DHL
GELQAM	D2	B	dynamic, with dispersion	DIV/DHV
WAKWA	D2	B	dynamic, no dispersion	DHV
RIWA	D2	B	steady state, no dispersion, PC-version	DHV
PROCNAAL	D2	B	dynamic	DHV
DIRTY	BD2		dynamic, with dispersion	Hmij
1-D horizontal models (streams)				
MODQUAL	D2	B	steady state, dispersion, dev.from QUAL-II	DHL
DIVAN	D1(2)	B	dynamic, no dispersion	DHL
DYNAQUAL	D2	B	dynamic, no dispersion	DHL
ESKWA	D2	B	steady state, dispersion	DHV
2-D horizontal models (seas, estuaries, large rivers, lakes)				
WAQUA	ABD12(4)		dynamic, reactions free as long as linear	DIV/DHL
SMOSS	D	B	distribution of oil slicks	TUD
3-D MODELS (seas, estuaries lakes)				
DELWAQ	D12(3)	AB	generic: free choice of grid, dev. from WASP	DHL
1-D vertical or layers models (stratified lakes and streams)				
OXY	E2	BC	for newly constructed reservoirs	DHL
ZQUA	E2	DBC		DHL
Zero-D or segment models (lakes, rivers)				
CHARON	D3(216)	B(ACE4)	focus on chemistry	
BLOOM-II 4	D		sequence of steady states, 10 algae species	DHL
BBSED	(FD)3(2)	(B)	focus on (sediment) chemistry	UT
MIDAS	4	B	diatoms and greens, variable cell stochiom	DHL
EXAMS	6(5F)	B	steady, toxics in solution, sed. organ.	EPA

[1] See figure 3 for A-F and 1-6 definitions
[2] DHL: Delft Hydraulics Laboratory; DHV: DHV Consultants, Amersfoort; DIV: Information processing Division Dutch Rijkswaterstaat; TUD: Technical University Delft; UT: University of Twente, Enschede; Hmij: Heidemij Consultants, Amersfoort.)

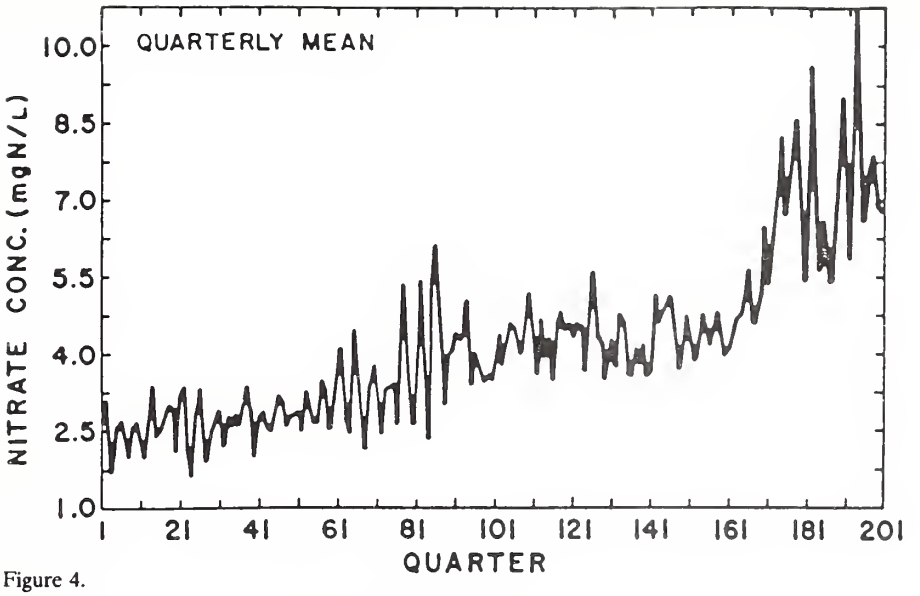


Figure 4.
Trends in Land Use change in the Thames Basin.

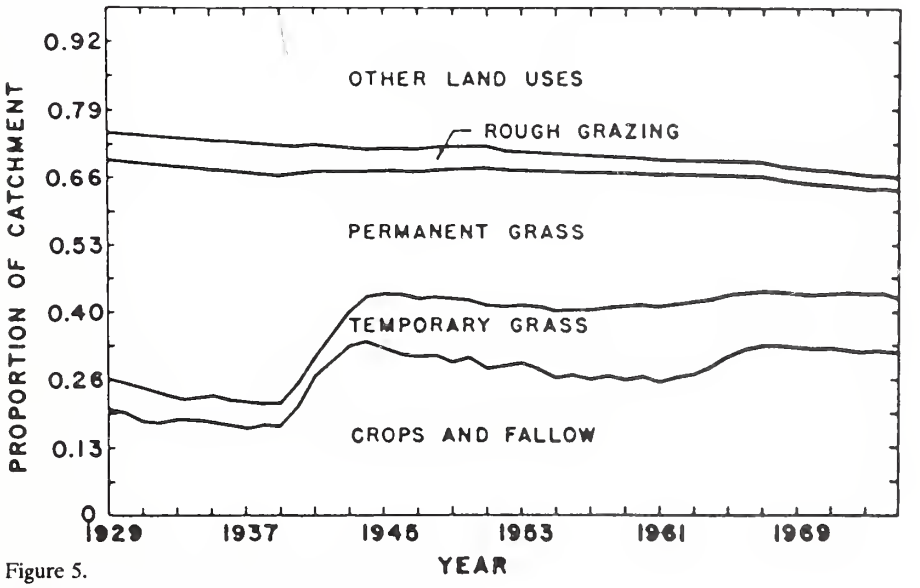


Figure 5.
Nitrate concentrations in the Thames at Walton.

In order to evaluate future trends and management options a series of component models for the river basin were developed. These included:

(1) a daily water quantity model for the Thames system which includes 17 tributary subcatchments and several major aquifer systems. The model provides river input flows such as tributaries, groundwater, surface runoff, effluent returns, and abstraction flows;

(2) a soil zone and aquifer model developed by the Water Research Centre for calculating the nitrate concentrations of surface runoff and groundwater given a particular landuse and fertilizer application rate;

(3) an integrated model of flow and water quality for the main river developed by the Institute of Hydrology. The model provides a mass balance along 22 reaches of the main river (see Figure 6), allows for denitrification processes and incorporates all inputs from the nonpoint sources derived by models (a) and (b) above. Details of the model are given by Whitehead and Williams (1984).

The integrated model has been used in a collaborative study with Thames Water Authority to investigate future nitrate concentrations in the Thames. It has been run using 56 years of hydrologic data to reproduce river flows and nitrate concentrations at Farmoor and Datchet. These abstraction sites provide drinking water for Oxford and London respectively. The model was run twice with two sets of conditions:

(a) 1982 agricultural landuse, fertilizer application rates, population levels and water demands, and
(b) 2006 agricultural landuse, fertilizer application rates. population levels and water demands.

The 1982 conditions were based on published land-use information and fertilizer surveys, together with inputs from sewage treatment plants based on historical population and per capita consumption data. Forecasts for the 2006 conditions were made from trends in the historical data and estimates of population growth.

The results from the models run under the two sets of conditions are summarized in table 5 which show the number of years the WHO limit of 11.3 mg N/l would be exceeded for various durations at the two abstractions sites considered. It is clear that at both abstraction sites there will be a marked increase in the number of years that this threshold would be exceeded if nothing were done to control the nitrate problem. The total number of years in which exceedance is predicted in any of the duration periods is 12 for Oxford and 7 for London for the 1982 scenario and 20 for London and 22 for Oxford for the 2006 scenario.

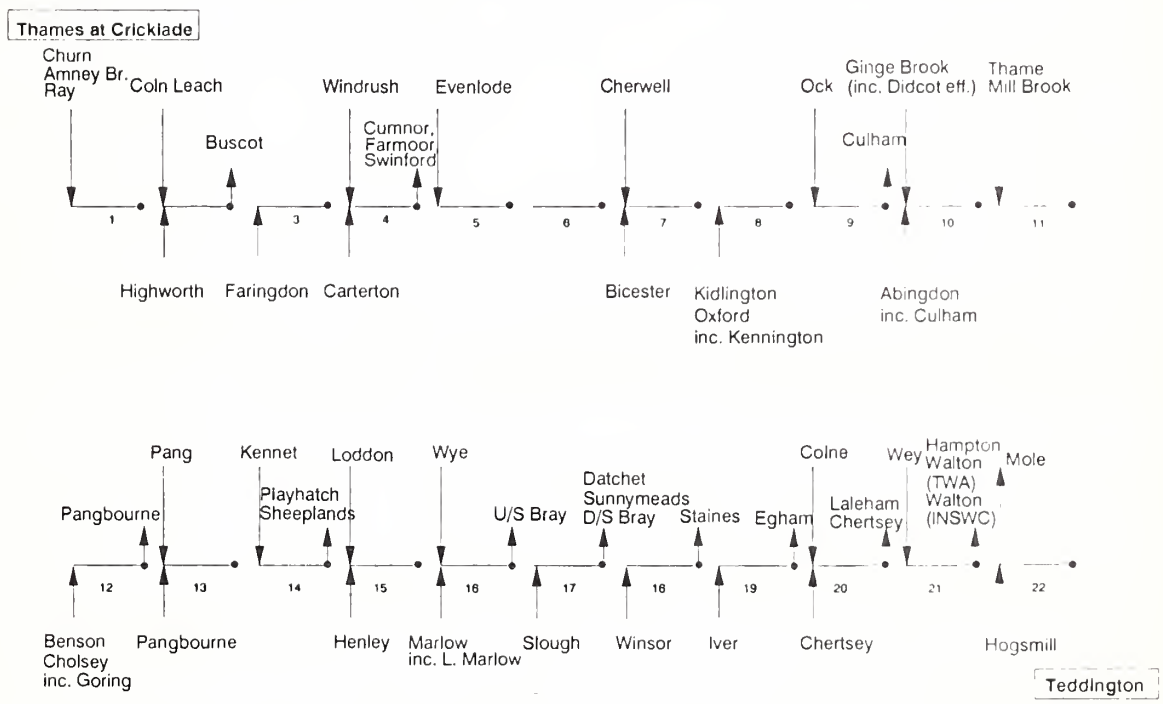


Figure 6.
Reach structure for Thames multi-reach model.

Table 5.
Number of years the 11.3 mg-N/l-threshold is exceeded in the October-September water year at London and Oxford on River Thames. Scenarios consist of landuse, population, and water demand.

Duration of exceedance (days)	Oxford		London	
	1982 scenario	2006 scenario	1986 scenario	2006 scenario
1-15	7	10	5	8
16-30	3	6	2	9
31-45	2	2	0	2
46-60	0	1	0	1
61-75	0	3	0	0
>75	0	0	0	0
as annual average	0	0	0	0

The model also provides information on the relative effects of diffuse pollution. During autumn storm conditions and winter high flows the nitrate diffuse sources such as agricultural runoff greatly exceeded the point-source pollution. The situation is altered during low flow conditions when point sources are relatively more significant.

The Lake Balaton Eutrophication Study

Lake Balaton (surface area 596 km^2, average depth 3.2 m, watershed area 5776 km^2) forms the most important recreational area in Hungary and its region is subject to intensive agricultural development. The lake has shown the unfavorable signs of human-induced eutrophication during past decades. By the early 1980's the water had reached a eutrophic- hypertrophic state calling for urgent short-term action. Detailed 1975-1981 observations indicated that the contributions of sewage were 28% and 52% of the total P (TP) and biologically-available P loads (865 kg/d and 465 kg/d, respectively). The ratio is opposite in character for nonpoint sources, 47% and 33%, respectively, suggesting that from the viewpoint of short-term management the sewage load is the dominant factor. For long-term planning the Zala River (average streamflow is around 9 m^3/s) contributes 25-30% of the lake's total load and leads to 13 times higher volumetric TP load at the most Western basin of the lake as compared to the less-polluted Eastern one.

The biology, chemistry, hydrodynamics, hydrology and meteorology of Lake Balaton have been intensively studied during the past 40 years. During the late 1970's and early 1980's an analytical systems approach was applied within the framework of international cooperation (with the leading role taken by the International Institute for Applied Systems Analysis (IIASA) the objectives were to understand procedures and evaluate feasible short-term eutrophication control strategies. The particular approach selected was based on the principle of decomposition and aggregation, and a model system following the structure of figure 1 was developed. Details of the case study can be found in Somlyody and van Straten (1986). Here we stress major points of the analysis from the viewpoint of applying models for planning and management purposes:

(1) though the validation of various dynamic ecosystem models was only partially successful at best, they gave the same (or similar) response to external P load

reductions, offering the practical conclusion that they can be used for planning purposes;

(2) an analysis has shown that the model system is sensitive only to weekly-monthly changes of external factors thus allowing the development of synthetic time series generators (based on available historical data) for forcing functions on corresponding time scales, and the usage of ecological models in a Monte Carlo fashion. This is a critical point of the analysis, since for shallow-lake, eutrophication-control problems no deterministic "critical" design scenario can be found (such as low-flow conditions for river DO problems) and inputs should be basically considered stochastic according to their nature;

(3) Not only an input generator but also an aggregated external- load scenario-generator expressing major control measures (e.g. tertiary treatment and establishment of pre-reservoirs) was coupled to ecological models. The result was an aggregated, planning- type stochastic load response model expressing the annual peak value of chlorophyll-a concentration (and its statistical variability), selected as indicator for management as a function of the total annual biologically available external P load and its variability (see later), as well as random variations in non-controllable meteorologic factors (such as solar radiation and temperature).

(4) A planning type nutrient load model was developed for the watershed. No detailed nonpoint-source modeling was involved, since realistic major control options were related to sewage treatment and diversion, and the construction of pre-impoundments at mouth sections of rivers before they enter the lake (lack of data, time and institutional difficulties also played roles in this respect). Such reservoirs can be well suited to the removal of phosphorus of both point- and nonpoint-source origins (through processes such as sedimentation, sorption, algal uptake, and benthic eutrophication etc). The basis for the derivation of the stochastic component of the load model was the daily observations initiated in 1975 at two cross-sections of the Zala River. Descriptive time series and multivariate-regression models (including various load components, precipitation streamflow, etc.) were developed and compared (Beck, 1982 and Somlyody 1986b). Time-series models were operated on a daily time scale, and regression models monthly, according to the requirement of the ecosystem model. Time-series models were capable of describing fast temporal changes induced by rainfall-runoff events; however, they did not capture the influence of slower variations of the hydrologic regime. Regression models (the development of which employed data for eight years aggregated a priori to monthly averages) were successfully calibrated and validated.

Eventually, for the synthetic generation of nutrient loads the regression model was selected and extended for the entire watershed. The uncertainty stemming from infrequent sampling (which was the case for all the smaller tributaries) was also analyzed (Somlyody et al 1986b) on the example of River Zala and subsequently extrapolated to the rest of the catchment;

(5) models already outlined were incorporated into a eutrophication optimization management model. Objective functions were formulated in terms of water quality goals expressed by maximum chlorophyll-a concentrations varying from basin to basin. Several mathematical models were employed, including a true stochastic approach an expectation-variance method and a deterministic procedure

(Somlyody and Wets 1985). Neglecting stochastic changes in external factors resulted in misleading the user to a basically incorrect strategy.

(6) The "optimal" short-term strategy worked out was basically accepted in the course of a year-long governmental policy-making procedure in 1982-1983 (in Somlyody and van Straten, 1986). Since 1983, phosphorus precipitation was introduced at ten treatment plants in the watershed: the regional sewage treatment system was extended; and the first element of the largest pre- reservoir (surface area of about 20 km^2), Small-Balaton at the River Zala, was put into operation in 1985 (its P removal efficiency is presently 50- 60%).

The management action plan schedules the achievement of the final (oligotrophic) water quality target level (corresponding to the lake's quality of the early 1960's) realistically for about 2005-2010, recognizing the time- lag effect caused by phosphorus accumulated in the sediment earlier (internal load) and the similar influence of "reconstructing" the present phytoplankton structure dominated by nitrogen-fixing, blue-green algae in summer periods.

Sample Cases from The Netherlands

A brief and fairly random selection of examples of model applications for management decisions in The Netherlands is given below. The examples have been selected to demonstrate:

(1) the use of models for planning and management,
(2) planning and management without the use of models,
(3) the use of a model for identification of research needs prior to management action,
(4) use of models for off-line operation, and
(5) use of models for on-line operational decision support.

Policy Analysis of Water Management in The Netherlands

This study, usually referred to as PAWN, was a cooperation between RAND corporation, the Delft Hydraulics Laboratory and the Dutch Rijkswaterstaat. Its aim was to analyze required national water management actions in The Netherlands to meet future demands (Pulles 1985). Although the key issue was related to surface water availability and water distribution, quality aspects played a part as well. Sea-salt intrusion water quality changes by distribution of River Rhine water, and eutrophication abatement were investigated. Over 40 models were used during the study. Restricting the discussion to eutrophication, the DHL-models BLOOM and CHARON were used to study effects of alternative management strategies on expected biomass levels in the Dutch lakes. These studies showed that total phosphorus concentrations would have to go to 0.1 mg/l in order to bring the algal biomass down to levels corresponding to chlorophyll-a of 50-100g/l in the majority of Dutch lakes. These levels are considered acceptable in the shallow Dutch lakes. Although the models were not able to compute the necessary load reduction to achieve these levels, it can safely be assumed that this corresponds to large reductions to some 10-20% of the present load (compare this to the target 50% P-load reduction in the Rhine River Action Plan agreed upon in the 8th Minister Conference of Rhine Shore States, 1 October 1987). Model results also indicated that flushing with nutrient-rich water (reducing the effective residence time available for algal growth) is hardly effective (see below for a counter case). The most effective would be deepening of the shallow lakes from 1-2 meters to 4-5 meters. No clear indication was obtained about the possible effects of dredging, since no reliable sediment models existed. The result of the modeling exercises largely contributed to the notion that only a concerted plan of action of several simultaneous measures could bring substantial improvement.

Lake Veluwe Eutrophication. Lake Veluwe is a case where models played only a remote role in the management decisions taken. Lake Veluwe is one of the artificial zone-lakes between the old land and the new Flevo-polder reclaimed from the former Zuiderzee. Some 15 years after the lake's creation in 1965, the water turned from clear to turbid, and extensive blooms of the blue-green algae Oscillatoria Aghardii have been observed ever since. In 1976 an extensive year-long program was started, aimed at management-oriented research and implementation of phosphorus-reduction measures, which was seen as one of the feasible approaches for restoration. Another option put into practice since autumn 1979 was to flush the lake with phosphorus-poor, calcium-rich polderwater from the Flevoland-polder during winter. The P-reduction and flushing decisions were not based on computer models for the lake, but were derived mainly from mass balance considerations (Projectgroep Eutrofieringsonderzoek Randmeren, 1986).

The flushing in combination with P-load reduction by tertiary treatment proved to be successful in reducing the total P levels in the lake. Also the biomass composition changed in favor of more green algae, and the almost exclusive dominance of the Oscillatoria was broken. The interesting aspect is that later model calculations with simple models using phosphorus apparent settling rates of P as a bulk parameter showed that the improvement in the lake's phytoplankton could not reasonably have been expected from flushing. Since the lake did, in fact, react, the models must have overlooked an important aspect. Sediment research suggested that possibly increased calcium- precipitation from calcium-rich flushing water created a larger adsorption potential in the sediments, thus reducing the release of P from the sediments (Van Straten 1986, Brinkman and Van Raaphorst 1986).

Poel en 't Zwet eutrophication. A phytoplankton-dynamics model was applied to the Poel en 't Zwet, a small shallow polder lake (Scholler et al. 1985). The internal load in this lake in summer turned out to be larger than the external load. Accordingly, model simulations showed that given a five- or ten-fold reduction in release from the sediment a reduction of chlorophyll-a from 400 μg/1 to 100 μg/1 was attainable. The simulations confirmed the belief of the water authority that dredging could be an effective measure, but the key question (whether the postulated reductions in P-release from the sediment would really be achieved by dredging) was not answered. Sediment P-release experiments are now being conducted to answer this question.

The River Vecht stormwater-overflow reduction. The GELQAM model was calibrated to DO and BOD data collected in the Vecht River during a stormwater- overflow event. The calibrated parameters were then used in an easy-to-use simplified PC-model (in Pascal) to assess the DO effects of overflows of stormwater of a specific size (in mm) and duration. The simplification was possible because the duration of stormwater events is limited. Therefore, it was sufficient to follow a train of water cells filled with overflow sewage on its way downstream. Dispersion effects were maintained. The required input, however, was largely reduced, and the desktop use in contrast to the mainframe GELQAM was an essential advantage. Each simulation represented a minimum DO point in a rainfall duration plot. Connecting the points of the lowest allowable DO yielded a line (the so-called isoDOx) which separated rainfall events that gave rise to DO-problems in the river from those that did not. The procedure was repeated for different management options, revealing that a reduction of critical DO-events in the Vecht River to no more than 4 times a year would require both the expansion of the municipal sewage treatment plant, as well as the reconstruction of the sewerage system of the city of Utrecht (Van Straten et al. 1986).

COMMON PROBLEMS WITH AVAILABLE MODELS: A USER'S PERSPECTIVE

The common problems arise as follows:

(1) **System Complexity** - river basin behavior is often extremely complex and it is necessary to identify the major variables and processes of interest and quantify these through field and laboratory studies. Model limitations need to be defined including biases and accuracy. Applications should not push the model beyond its original designed use or the level of detail incorporated into model processes.

(2) **Generalized Models** - often generalized models are applied with insufficient understanding of the particular river basin. There can be problems in applying a generalized model to a river system and it may be better to derive a model to suit the river system.

(3) **Critical Loads** - 90% of nonpoint pollution is often carried into river systems during the 10% or fewer events high flow. When building models to predict non point pollution the importance of hydrologic extremes must be considered.

(4) **Inadequate Calibration and Validation** - it is essential that models must be calibrated and validated using different data sets. It is important that reasonable ranges of parameters are identified and some assessment of parametric uncertainty on model predictions is made.

(5) **Poor Databases** - often models are developed and driven using poor data. It is essential that adequate reliable data (including all relevant processes) on river systems are collected.

(6) **Misinterpretation of Results** - It is fairly common for managers to misinterpret model results. Care must be taken in applying models, and it is essential to investigate how the uncertainties in parameters and input data affect model predictions.

(7) **Field Experiments** - Because of the lack of models and modeling knowledge decision-makers often resort to costly field experiments. Too often the relatively small additional investment to evaluate the experimental results with models is not made. This is almost always a false economy, because model implementation can steer experiments and save costs, especially if the modeling is done in parallel with the experiment. Such model application would assume that rapidly implementable tools are available, producing output that can be understood by the experimenters. Moreover, in the end, the modeling is a valuable permanent product of the experiment, applicable in other situations.

FUTURE MODEL REQUIREMENTS

Model development requirements can be specified on the basis of the state of existing models and pollution problems which are likely to emerge in the future. A list far from complete is as follows:

(1) development of better based watershed nonpoint-source models primarily for planning;
(2) enhancement of regional modelling for planning and management;
(3) models for supporting policy making (multi-objective methods, decision support systems etc.);

(4) involvement of self-organizing mechanisms into water ecosystem models. Modeling the behavior of blue-green algae and the sediment (especially in shallow water bodies);

(5) techniques of parameter estimation, calibration, identification etc. well suited for handling structural models with relatively large number of state variables and parameters;

(6) study of ill-defined systems (and issues such as uniqueness and stability).

There is a strong need for easily accessible computer models (preferably on PC's) for application to problems of nonpoint-source pollution.

There is a need for easily accessible development software that enables the user to develop and test new dynamic model concepts (generic software), without the need to reconsider the problems of numerical integration, data handling and plotting.

More attention is needed on the "input" side of the models: the use of models to generate results for planning purposes should take into account the stochastic nature of future inputs. Statistical data analysis packages are widely available, but generic tools for synthetic data generation for probabilistic simulation are still rare.

CONCLUSIONS

There is an increasing problem of nonpoint-source pollution derived from agricultural sources in the U.K. and continental Europe. Nutrients such as N and P from agricultural sources increase the eutrophic state of rivers and reservoirs. Pesticide usage is also increasing at an alarming rate.

Models of agricultural pollutants are at an early stage of development and planning design and operational models are required to adequately manage river basins. Generally water authority managers are willing to use models but often have to rely on outside experts to advise them on model use. There is a need for models that can be run on personal computers so that managers can become more familiar with model application and use.

REFERENCES

Beck, M.B. 1982. Time-series Analysis of Zala River Nutrient Loadings, Working Paper WO-82-116 (Laxenburg, Austria: International Institute for Applied Systems Analysis).

Beck, M.B. 1985. Water quality management: A review of the development and application of mathematical models; Lecture Notes in Engineering Series, 11, Springer-Verlag, 107 pp.

Beck, M.B. ed. 1987. Systems Analysis in Water Quality Management. Proceedings of a Symposium held in London, UK, 30 June - 2 July 1987. Pergamon Press, Oxford-New York-Beijing-Frankfurt-Sao Paulo-Sydney-Tokyo-Toronto.

Benndorf, J., and Recknagel, F. 1982. Problems of Application of the Ecological Model SALMO to Lakes and Reservoirs having Various Trophic States. Ecol. Model., 17.

Bogardi, I., David, L., and Duckstein, L. 1983. Trade-offs Between Cost and Effectiveness of Control of Nutrient Loading into a Water Body. Research Report RR-83-19 (Laxenburg, Austria: International Institute for Applied Systems Analysis).

Bogardi, I., and Duckstein, L. 1978. Input for a stochastic control model of P loading. Ecol. Model. 4.

Brinkman, A.G., and Van Raaphorst, W. 1986. De fosfaathuishouding in het Veluwemeer (The phosphorus household in Lake Veluwe: dissertation, University of Twente, 481 pp. annex, [in Dutch].

Brown, L.C., and Barnwell Jr., Y.O. 1986. The Enhanced Stream Water Quality Models QUAL2e and QUAL2e-UNCAS: Documentation and User Manual. Environmental Research Laboratory Office of Research and Development US Environmental Protection Agency, Athens, Georgia. EPA/600/3-87/007, May 1987.

Crabtree, R.W. 1986. River Water Quality Modelling in the UK, WRC Report ER 196E, UK.

Hornberger, G.M., and Spear, R.C. 1980. Eutrophication in Peel Inlet, I: Problem defining behavior and a mathematical model for the phosphorus scenario. Water Res. 14.

Jolankai, G., and Pinter, J. 1886. Water quality and management modelling of non-point source pollution - a case study. IAHS Publ. No 157.

Kozerski, H-P., and Dvorakova, M. 1982. A Three-Layer Aquatic Ecosystem Model: Sediment Submodel Tests and Some Simulations. Ecol. Model. 17.

Leonov, A. 1985. Modeling and Explaining the Phosphorus Dynamics of Lake Balaton, 1976-1979. Research Report RR853. (Laxenburg Austria: International Institute for Applied Systems Analysis).

Onstad, C.A., and Blake, J. 1980. Thames Basin and Agricultural Relations. Proceedings Symposium on Watershed Management ASCE/Boise.

Pinter, Gy., and Feher, J. 1987. Management Planning Tools to Control Non- point Source Pollution. Seminar on Protection of Soil and Aquifers Against Non-point Source Pollution, Madrid (Spain) 5-9 October 1987.

Projectgroep Eutrofieringsonderzoek Randmeren. 1986. Bestrijding van de eutrofiering van het Veluwemeer (Abatement of eutrophication of Lakes Veluwe and Dronten), Rijkswaterstaat - DBW, Lelystad, 296 pp.

Pulles, J.W. 1985. Beleidsanalyse voor de waterhuishouding in Nederland (Policy Analysis of Water Management for the Netherlands), Dutch Rijkswaterstaat, The Hague, 431 pp. [in Dutch].

Recknagel, F., Benndorf, J. and Uhlmann, D. 1986. Model-Aided Water Quality Control of Standing Waters. Proceedings of the International UNESCO Symposium on Decision Making in Water Resources Planning, Oslo, Norway, May 5- 7, 1986.

Scavia, D., and Robertson, A. 1979. Perspectives on Lake Ecosystem Modeling, Ann Arbor, MI: Ann Arbor Science Publishers.

Scholler, M., Van Straten, G., and Fieggen, W. 1985. Eutrofieringsmodel ten behoeve van het waterkwaliteitsbeheer in de meren van be polder Wormer, Jisp en Neck (Eutrophication model for water quality management in the lakes of the polder Wormer, Jisp and Neck); H20, 18: 548-554. [in Dutch].

Somlyody, L. 1978b. Mixing Model of Large Streams. Progress in Water Technology, vol. 10. Nos. 5/6.

Somlyody, L. 1982. Water quality modelling: A comparison of transport- oriented and ecology-oriented approaches, Ecol. Model. 17.

Somlyody, L. 1986a. Wind induced sediment resuspension in shallow lakes. Proceedings of the Int. Conference on Water Quality Modelling in the Inland Natural Environment, Bournemouth.

Somlyody, L. 1986b. Nutrient loads of rivers: time series and regression models. In Monitoring to detect changes in water quality series. Proceedings of the Budapest Symposium, July 1986. IAHS Publ. No. 157.

Somlyody, L., and van Straten, G. eds. 1986. Modeling and Managing Shallow Lake Eutrophication. With Application to Lake Balaton. Springer Verlag. Berlin-Heidelberg-New York-Tokyo.

Somlyody, L., and Wets, R. 1985. Stochastic Optimization Models for Lake Eutrophication Management. Collaborative Paper CP-85-16 (Laxenburg, Austria: International Institute for Applied Systems Analysis).

Steel, J.A. 1975. The management of Thames Valley Reservoirs Proceedings of Conference in the Effects of Storage in Water Quality, WRC, UK.

Straskraba, M. 1985. Managing of eutrophication by means of ecotechnology and modelling. Lakes Pollution and Recovery. Proc. Internal. Congress, Rome 15- 18th April 1985.

Straskraba, M. and Gnauck, A. 1985. Freshwater Ecosystems. Modelling and Simulation. Developments in Environmental 8. Elsevier, Amsterdam-Oxford-New York-Tokyo.

Svirezhev, J.M., Luckyanov, N.K., and Logofet, D.O. 1979. Simulation Model of an Oligotrophic Lake. In Jorgensen, S.E. (ed.). Hydrophysical and Ecological Models of shallow Lakes and Reservoirs. Summary Report of an IIASA Workshop, April 11-14, 1978. CP7814. International Institute for Applied Systems Analysis, Laxenburg, Austria.

Van Straten, G., Van Riel, P.H., and Heins, J.C. 1986. Model-based assessment of effects of storm water overflow management upon low dissolved oxygen occurrence in the receiving stream; In Urban Storm Water Quality and Effects Upon Receiving Waters, International Conference, Wageningen, The Netherlands (Committee on Hydrological Research, proceedings No. 36, The Hague).

Van Straten, G. 1986. Identification, Uncertainty Assessment and Prediction in Lake Eutrophication; dissertation, University of Twente, 240 pp.

Vavilin, V.A., Vasilyev, V.B., Tsitkin, M.Yu., and Bagotskii, S.V. 1979. The Prediction of Solution to Management and Engineering Problems Related to the Eutrophication of Water Bodies. In Jorgensen, S.V. (Ed.). Hydrophysical and Ecological Models of Shallow Lakes and Reservoirs. Summary Report of an IIASA Workshop, April 11-14, 1978. CP-78-14. International Institute for Applied Systems Analysis, Laxenburg, Austria.

Whitehead, P.G., Young, P.C., and Hornberger, G.E. 1979. A systems model of flow and water quality in the Bedford Ouse River system: Part I, Streamflow Modelling, Water Res. 15: 1157-1171.

Whitehead, P.G., Beck, M.B., and O'Connell, P.E. 1981. A system model of flow and water quality in the Bedford Ouse River system: Part II, Water Quality Modelling, Water Res. 15: 1157-1171.

Whitehead, P.G., Caddy, D.E., and Templeman, R.F. 1984. An on-line monitoring data management and forecasting system for the Bedford Ouse River Basin, Wat.Sci. and Tech. 16: 295-314.

Whitehead, P.G., and Hornberger, G.E. 1984. Modelling algal behavior in the River Thames, Water Res. 18: 945-953.

Whitehead, P.G., and Williams, R. 1984. Modelling nitrate and algal behavior in the River Thames, Wat.Sci.and Tech. 16, pp 621-633.

PROBLEMS WITH SURFACE WATER MODELS FROM A USER'S PERSPECTIVE

Kent Thornton[1], Claire Stalnaker[2], and Kenneth Baun[3]

ABSTRACT

The model user has not been an integral part of, or, in many cases, considered during model development. Greater model utility can be achieved by providing: 1) estimates of prediction uncertainty, 2) model documentation and user manual development, 3) model support groups - either state or federal, 4) international user's workshops for evaluating new and modified models, and 5) post-audit analyses to determine the accuracy of previous model applications.

INTRODUCTION

The role of surface water models in water resources management has increased dramatically over the past two decades. Few water resources issues are resolved without the use and application of surface water models ranging from relatively simple empirical models to complex, process-oriented basin models. The models reflect the complexity of the water resources issues ranging from minimum stream flows for water supply, irrigation, fish and wildlife habitat, water quality, recreation, and aesthetics to basin-wide wasteload allocation or eutrophication studies to assessing the fate, transport, and transformation of pollutants and hazardous substances in our surface waters. Multiple desired water-resources uses result in multiple, interactive water resources issues that must be resolved. Surface water models provide a flexible, efficient, and cost-effective approach for integrating many process interactions and evaluating the effects of various management alternatives on surface water quantity and quality.

Surface water models have been used and applied to aquatic problems for over 50 years. The classic studies of Streeter and Phelps (1925) on the biochemical oxygen demand (BOD)-dissolved oxygen relationships in the Ohio River form the basis of many wasteload allocation models currently in use, particularly desktop modeling approaches. Model development and refinement has incorporated additional processes and additional spatial and temporal dimensions from simple steady-state point models to three-dimensional, time-varying models incorporating hydrodynamic, chemical, and biological kinetic formulations. The introduction and refinement of personal computers (PC's) has provided model developers and users with tremendous computing power unavailable five years ago. The efficacy of using surface-water models on complex water resources issues, coupled with increased computing capabilities has spurred the development of a greater number of sophisticated surface water models. An U.S. Congress Office of Technology and Assessment (OTA) survey, for example, indicated direct and indirect federal expenditures for model development and dissemination during 1979 amounted to about $50 million (Friedman et al. 1984). Increased model development and dissemination, however, are not paralleled by greater ease or efficiency in model usage. In fact, the correlation between model proliferation and model documentation and support might be negative. It is this concern over the relation between model development and effective model utilization by federal, state, and private users that is the topic of this paper.

[1]Kent Thornton, FTN Associates, Ltd., Little Rock, AR.

[2]Claire Stalnaker, USDI Fish and Wildlife Service, Ft. Collins, CO.

[3]Kenneth Baun, Wisconsin Dept. of Natural Resources, Madison, WI.

This paper will consider five areas that should be addressed by the modeling community to improve the use and application of surface water models. These areas are:

(1) prediction uncertainty,
(2) documentation and manual development,
(3) model support,
(4) user's workshops, and
(5) post-audit analyses.

These five topics represent generic problems for users and reflect the private, federal, and state perspective of the authors, respectively. While these problems are not new to model users, they are still largely unresolved and, therefore, warrant further discussion.

PREDICTION UNCERTAINTY

Phrases such as "adequately describes" or "captures the essence", in general, are used in the modeling literature to describe the relation between model predictions and surface water observations. For most model applications, however, it is critical to quantify the uncertainty in the predictions so realistic comparisons can be made among management alternatives. Model predictions generally are assumed to represent the expected system response, even though model predictions have not been confirmed or validated. A recent survey was conducted among hydrologic modelers, who were asked to assess the accuracy of models in predicting the effects of major landuse changes on the hydrology of sites with no site calibration data and limited validation data. Most of the 28 hydrologic models were judged capable of providing good accuracy (Task Committee 1985). The reasons for this rating appeared to be based on personal experience, technical knowledge of assumptions, limitations and applications, and belief in the model originators (Task Committee 1985). Comparisons of model predictions based on hypothetical catchment data for five of the 28 models, however, did not support this confidence in model accuracy. For three discrete-event, continuous-process models, the maximum differences among model simulations were 139% for runoff volume, 69% for peak discharge and 37% for time to peak (Task Committee 1985). Two continuous watershed models differed by a factor of six in estimating annual runoff volumes (Task Committee 1985). Verification by reputation is unsatisfactory when differences among various management alternatives might represent thousands to hundreds of thousands of dollars. It is critical that users have an estimate, or be able to estimate, prediction uncertainty and perform statistical comparisons among management alternatives. Implementing a management practice without knowing the uncertainty in the model predictions can be both environmentally and monetarily costly. Multilevel or selective-withdrawal outlets in reservoirs, for example, provide greater operational flexibility in meeting downstream water quality and environmental standards and criteria. These outlet structures, however, add several million dollars to the cost of the project. Before these expenditures can be justified, the decision-maker must be assured the environmental conditions will be significantly improved with the selective withdrawal structure. Without some estimate of prediction uncertainty, this decision can be difficult. We modified a deterministic reservoir water quality model to incorporate data and prediction uncertainty and used this model to evaluate different withdrawal alternatives in meeting in-lake and downstream water quality (Thornton et al. 1983). Differences in release water quality were noted with the deterministic simulations (fig. 1a) but it was not known if these differences were sufficient to warrant a $5-10 million selective withdrawal structure. Stochastic simulations, incorporating sampling, analytical and parameter uncertainty, clearly indicated there was no statistically significant difference among withdrawal alternatives and that the less expensive bottom withdrawal structure was adequate (fig. 1b). The decision-maker must have this type of information to make cost-effective, environmentally sound decisions.

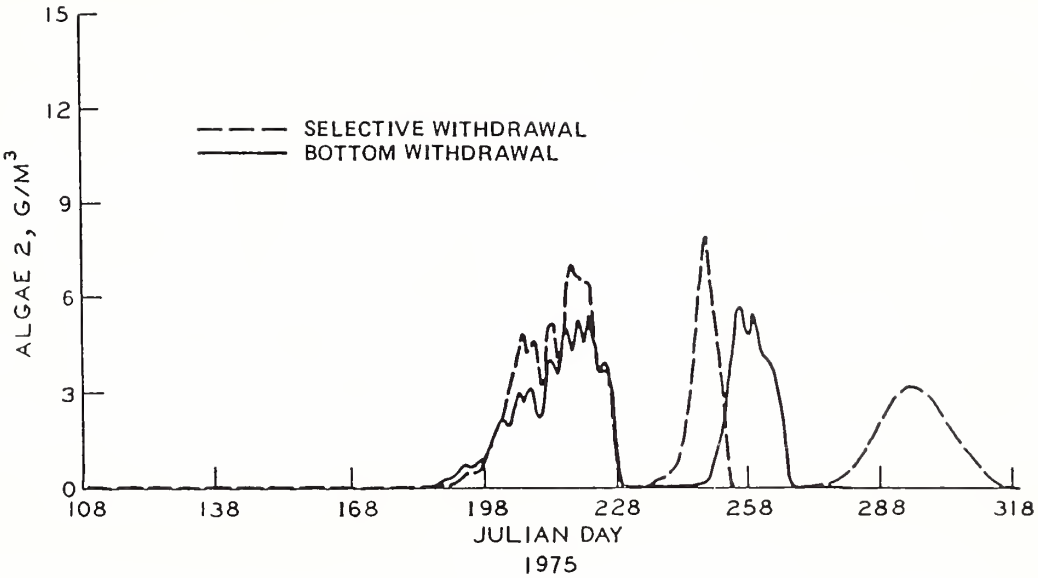


Figure 1a.
A deterministic simulation of two reservoir withdrawal alternatives. It is not clear if these trajectories are different.

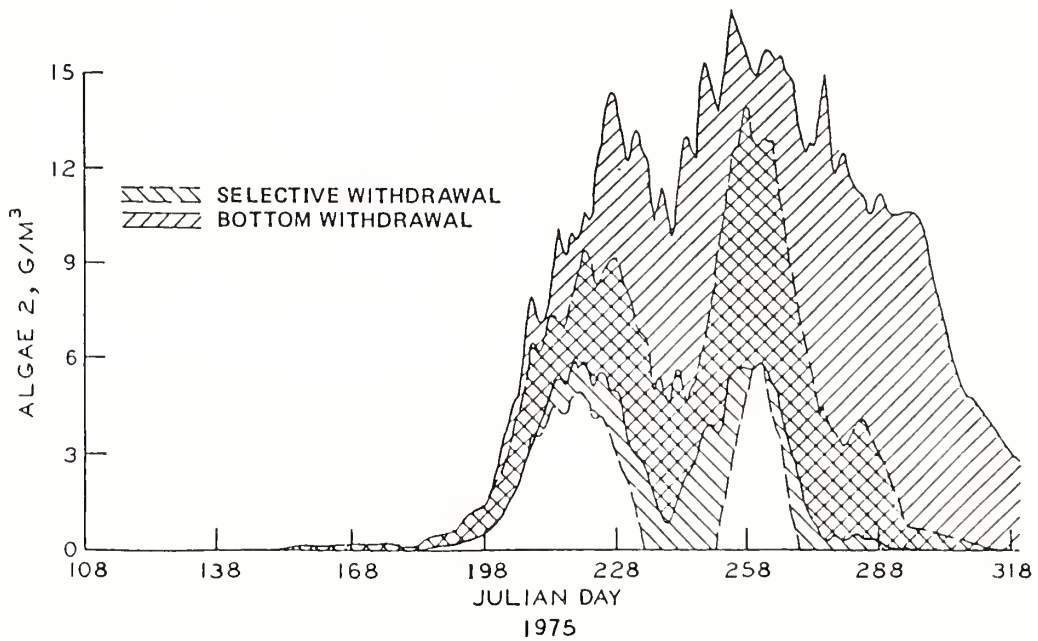


Figure 1b.
Stochastic simulations indicate the two alternatives are not significantly different. The difference in design and construction costs between these two alternatives was about $5 million.

Procedures for uncertainty and error estimation in surface water modeling have been described by Spear and Hornberger (1980), Thomann (1982), Reckhow and Chapra (1983), and Beck (1987). These procedures include a variety of techniques such as first order error analysis, regional sensitivity analysis, Monte Carlo simulations, methods of moments and maximum likelihood estimators. These procedures, although not discussed here, permit an evaluation of prediction uncertainty. Model developers typically focus on minimizing prediction uncertainty and increasing model resolution. While this is obviously desirable, quantifying the prediction uncertainty can be more important. For specific user needs, the existing model resolution might be satisfactory and sufficient to assess various management alternatives. Many agencies are focusing on risk assessment and identifying the risks associated with selected decisions. The Environmental Protection Agency (EPA), for example, is developing approaches for risk characterization, risk assessment, and risk management based on potential impacts and underlying uncertainty (Thomas 1987). Quantifying prediction uncertainty is an important component of risk characterization. Model users must be able to assess prediction uncertainty, which should be a critical research area for model developers.

DOCUMENTATION AND MANUAL DEVELOPMENT

Model documentation and well-written, informative user's manuals are glaringly absent for many models. Journal articles, written for peers, on model theory, are neither satisfactory documentation nor suitable for manual development. Without adequate documentation, the underlying assumptions, limitations, process formulations, and time and space scales inherent in the model are unknown to the user. This will result in inappropriate applications of the model. Inappropriate application also can occur when user's manuals are not available. A user's manual must accompany the distribution of every model. Information on model language, operating systems, flow charts, input requirements, format and preparation, and output format and examples are critical for proper model usage. Calculations using nitrate (NO_3) or phosphate (PO_4) concentrations for example, rather than the elemental NO_3-N or PO_4-P concentrations can have profound effects on predicted algae concentrations and grossly overestimate required watershed best management practices or in-lake techniques to reduce nutrient loads.

There are obvious exceptions, such as the Corps of Engineers' Hydraulic Engineering Center (HEC) manuals, and EPA's Stormwater Management Model (SWMM) Extended Transport (EXTRAN) manual, which should serve as examples for model documentation and user's manuals.

The general organization of these user manuals is:

(1) Introduction and background (including identification of support group).
(2) Operating requirements.
(3) Input requirements.
    (a) overview
    (b) detail and preparation
(4) Example problem.
    (a) input format and preparation
    (b) output format
    (c) discussion
(5) Trouble-shooting tips.
(6) Appendices.
    (a) detailed theory
    (b) flow charts
    (c) software and hardware requirements
(7) References.

These chapters should, at a minimum, be included in every model user's manuals. The inclusion of chapters on example input and output for various simulations, general interpretation, and trouble-shooting tips are extremely useful for most users.

In application, model users tend to operate in a realm of limited staff resources. There simply isn't the staff available to spend large amounts of time preparing data or interpreting output. To be used then, a model must be practical. To be practical, a model must have reasonable data input requirements and must provide easily-understood data output.

Model developers can facilitate data preparation and output interpretation in several ways. They can:

(1) Guard against overparameterization,
(2) Tie the model input and output to a data base management system, and
(3) Provide associated software to help collect and check the input data, or tabulate, collate, or graph the output data.

Overparameterization is when a model requires data input that is not justified by the resulting increase in precision of the model output. A two-percent increase in model accuracy may not justify a ten percent increase in model input requirements. A balance between reasonable input and reasonable output must be struck. Comparisons among alternative formulations, their input requirements, and prediction accuracy can provide some of this information.

Data preparation is greatly enhanced if the model is able to read input files directly from a data base management system. Many models require ASCII data input files. While anyone with a simple ASCII text editor can build these files, they are relatively difficult to assemble and modify. A good data base management system will not only facilitate data entry and modification, but also will be able to report, tabulate, and graph the input and output data as well.

Lastly, the model really has to operate as a system in conjunction with other supportive software. If ASCII files are required, does the model come with software that will facilitate data entry? Or does it come with programs that will check the data before the model is run? What sort of associated programs are there for displaying or interpreting the model output? If these types of programs are not supplied with the model, it is likely the model user will have to create them.

There is a good likelihood that the model user may find it necessary or desirable to modify a model to suit their particular needs or situation. Model developers can facilitate model modification in several ways. They can:

(1) Provide the model source code along with the executable code,
(2) Provide adequate comments within the source code,
(3) Provide a data dictionary of variable names, definitions and units, and
(4) Ensure that the source code meets good programming standards and conventions.

While it may be difficult for a model developer to give up exclusive ownership in a program, in the long run it may mean that the model receives much better reception and application.

From a user's perspective, the user's manual <u>is</u> the model. The user's manual reflects the programming, accuracy, precision, level of detail, flexibility, ease of use, and applicability of the model. The user's manual, therefore, should be reviewed and edited as extensively and exhaustively as any refereed journal article. Functioning model code, running on several computer systems with quantified uncertainty estimates, represents the <u>first</u> step toward the development of a useful model, not the last step.

MODEL SUPPORT

The importance of model documentation and user's manuals cannot be understated, but these documents must be supplemented by adequate model support. Questions will continually arise that cannot be addressed without a support person or group. OTA conducted an analysis of federal and state agency use of water resources models in planning, management and policy. One of the major findings of this study was:

> "Successful modeling requires adequate resources for support services, such as user assistance, as well as for development. Presently, model development has outstripped corresponding support for models. In the past, model developers have put a premium on the development of models, while support for models - documentation, dissemination, user assistance, and maintenance - has been neglected. Often, resources are focused on development but are unavailable for support activities. The neglect of model support has led to a multiplicity of models, most of them under-utilized. Many of these models cannot be used by personnel other than the developer because of lack of documentation, access to the model, and user assistance."

Model support, therefore, is, and must be, an integral part of model use.

Model support, however, requires time, personnel, and money. There are support groups for some classes of surface-water models such as the HEC supporting hydrologic models, EPA's Athens, GA, Environmental Research Laboratory, support groups for water quality models and the Holcomb Research Institute in Indianapolis, IN supporting groundwater models. Additional funding and personnel are required if these agencies or institutes are to support other models. Similar support groups need to be established for various classes of models. These support groups could be funded by user fees and supplemented by state and federal funds. There still are, however, only a limited number of models that can realistically be supported by any group.

These supported models, however, can be a continually evolving group. As better models are formulated, documented, and used, some models should be discontinued and replaced. An older version of a model that is heavily used by a particular private, state, or federal agency probably requires minimal external support because of user familiarity with the model. The support group, therefore, can focus on learning, understanding and using the new models or newer versions of existing models so support can be provided to the user community. The new models need to be continually reviewed and tested to determine which models should be considered for addition to the support groups. This review and testing can occur through the establishment of specific evaluation criteria, user evaluation workshops and test centers.

USER WORKSHOPS

Workshops, symposia, and seminars are frequently used to present and discuss new models. International comparisons of models have been sponsored for hydrometeorological models to evaluate model accuracy with reference to actual and hypothetical data. Similar regional and international workshops for specific classes of surface-water models could be conducted, emphasizing model usage. Regional workshops can focus on the use, required modifications and accuracy of models for region-specific applications. International workshops can focus on the transportability, general utility and robustness of models for applications. These workshops should be held periodically (e.g., every 3-4 years) at or near a support group's facilities, and selected federal, state, and private users should be invited to participate in the workshop. The user's manual and documentation would be provided to the participants prior to attendance for perusal so actual "hands-on" usage would occur at the workshop. Locating discrepancies between

documentation and code, identifying deficiencies in manual descriptions and instructions, and rating the models using selected criteria (e.g., criteria such as in table 1) would occur during the workshop. This information would be compiled into a workshop summary and used to determine which, if any, models should be supported and which existing models would be replaced. This information, along with the compiled workshop model evaluations, would be published for review by the modeling community.

A beta-test center concept could be used to screen models during the 3-4 year interim period between workshops. A voluntary system could be established with individual or group users serving as interim reviewers for new models. These reviews could be compiled by the support group and used to screen models and eliminate those that clearly do not have widespread applicability. The periodic workshops, then, would evaluate and demonstrate those models that appear to have general applicability based on preliminary test center review. This review could be enhanced by periodic review of previous applications and post-audit studies.

Table 1. Criteria and rating of surface water models.

Criteria	Rating Range
User manual and documentation	Excellent.....Poor
Validation/application results	Excellent.....Poor
Temporal resolution	
Time scale	Minute---Annual
Steady state	Yes/No
Dynamic	Yes/No
Event-based	Yes/No
Continuous	Yes/No
Spatial resolution	
Point	Yes/No
Dimensionality	1-D.....3-D
Modules	
Hydrologic	Routing
Hydraulic	Steady Q.....Unsteady Q
Chemical	N, P, C.....Se
Biological	Phyto.....Fish
Input data requirements	Low.....Extensive
Output format	Graphic, Tabular
Uncertainty estimates	1°Error.....None
Computer requirements	FTN 77, 350K
Manpower requirements	Jr Engr.....Ext. Exper. Multiple People
Recommended applications	Habitat Modifications Nonpoint Source (NPS) Wasteload - small streams

POST-AUDIT STUDIES

Most model applications are used as part of planning and management studies such as wasteload allocation studies, watershed landuse management alternatives, reservoir design and operation, minimum low flow estimates and similar uses. The accuracy of these predictions following design and implementation, however, is rarely evaluated. Numerous environmental impact studies have been conducted throughout the country using a variety of modeling approaches. Post-audit studies of the forecasts and predictions generally are not available. These post-audit studies would benefit both model developers and model users.

The OTA could be charged with conducting post-audit studies to complement the OTA review of model usage conducted in 1982. In addition, post-audit studies should be factored into federal budgets for modeling activities. These studies could provide better information for both model development and model use than additional research into specific environmental processes and process formulations.

Systems with quality-assured, post-monitoring data would be excellent candidates for these post-audit studies. Case studies could be conducted for each of the general types of surface water models. These case studies should be published in the refereed literature, included with the support-group review documents and distributed with the user's manuals for these models.

CONCLUSIONS

Modeling is an iterative process. Model developers, model users, and decision makers must interact if models are to be effectively used in surface water planning, management and policy. The five areas discussed above can be used to effectively integrate model development with model usage and provide better products for decision makers.

REFERENCES

Beck, M. B. 1987. Water quality modeling: A review of the analysis of uncertainty. Water Resour. Res. 23:1393-1442.

Friedman, R., C. Ansell, and S. Diamond. 1984. The use of models for water resources management, planning and policy. Water Resour. Res. 20:793-802.

Mittleman, A., H. Shugart, R. Park, and others. 1987. Ecological model selective criteria. TRI Draft Report. United States Environmental Protection Agency, Washington, DC.

Reckhow, K. H., and S. C. Chapra. 1983. Engineering approaches for lake management. Ann Arbor Science, Boston, MA.

Spear, R. C., and G. M. Hornberger. 1980. Eutrophication in Peel Inlet, II. Identification of critical uncertainties via generalized sensitivity analysis. Water Resour. Res. 14:43-49.

Streeter, H. W., and E. B. Phelps. 1925. A study of the pollution and natural purification of the Ohio River III. Factors concerned in the phenomena of oxidation and reaeration. Public Health Bull. No. 146. U.S. Public Health Service.

Task Committee on Quantifying Land-Use Change Effects of the Watershed Management and Surface-Water Committees of the Irrigation and Drainage Division. 1985. Evaluation of

hydrologic models used to quantify major land-use change effects. ASCE Irrig. Drain. Eng. 111:1-17.

Thomann, R. V. 1982. Verification of water quality models. ASCE Environ. Eng. Div. 108:923-940.

Thomas, L. M. 1987. Environmental decision-making today. EPA Jour. 13:2-

Vollenweider, R. A. 1968. Scientific fundamentals of the eutrophication of lakes and flowing waters, with particular reference to nitrogen and phosphorus as factors in eutrophication. OECD. Paris Tech. Rept. DA 5/SCI/68.27.

DISCUSSION OF PAPERS PRESENTED IN TECHNICAL SESSION 4, PART 2: SURFACE WATER MODELS FROM THE USER'S PERSPECTIVE

Rick Allen[1], Presiding
Jim Doughtery[2], Recorder

PAPERS DISCUSSED

Surface Water Quality Models for Planning, Design and Operational Management by P.G. Whitehead, L. Somlyody and G. Van Straten

Problems with Surface Water Models from a User's Perspective by K. Thornton, C. Stalnaker and K. Baun

SUMMARY

The discussion centered on two main topics. The first was model standardization, the second was the general utility of models.

With respect to model standardization, the major concern was documentation, validation and support of models. Poor model documentation was a significant concern to several commentators. A solution to this problem was suggested to be the withholding of a portion of the development money until the funding agency received what it considered to be acceptable documentation. The type of information that is contained in this documentation should be of two main types. The first type is technical support to model users. This information would include example runs and solution algorithms. The model user would use this information to make sure the model was operating correctly. The other type of information would be oriented toward the decision-maker. This information would include assumptions, limits of applicability of the model and appropriate as well as inappropriate applications of a particular model.

Additional support to model users was suggested in the form of governmental agencies that would establish support groups for models with a particular type of application. For example, the U.S. Geological Survey and the Environmental Protection Agency may join together to form a support organization for groundwater models. This would be in the agencies' best interest since these models would generally be used in reference to activities that the agencies are undertaking. This group would act as a clearinghouse that would produce standard models and provide support information as well as information on novel applications of particular models. This organization would also monitor the applications of these models and evaluate their utility and limitations.

Two examples of this approach were given. The International Association for Water Pollution Control has developed a state-of-the-art approach to model-activated sludge treatment processes. The association has developed an IBM-PC-compatible version of this model and provides workshops on use of the model. This approach has the advantages that information from all over the world has been collected in a standardized way and the model can be consistently evaluated on a large number of different data sets. Another example is that of the Hydrologic Engineering Center and its activities in flood-plane modeling.

[1]Rick Allen, Assistant Professor, Utah State University, Logan, Utah.
[2]Jim Doughtery, Department of Civil Engineering, Utah State University, Logan, Utah.

Related to this is the discussion of the applicability of these models in predicting what actually occurs in the field. The extreme simplifications of some of these models make the utility of their predictions very limited. Without a better understanding of the complex chemical, physical and biological processes, it may be futile and self-deceptive to progress into models of higher degrees of complexity.

The defense of models was that they provide a consistent and quantifiable means of allocating resources. Due to the large amount of litigation, models seem to be required so that decisions can be shown to have taken place within a consistent framework. This does not justify application of models in a vacuum. The models results must always be considered in light of previous experience and limitations of the model.

ENVIRONMENTAL ISOTOPE TRACER STUDIES OF CATCHMENT PROCESSES: TOOLS FOR TESTING INTEGRATED WATER QUALITY MODELS

Michael G. Sklash[1], Ian D. Moore[2], and Gordon J. Burch[3]

ABSTRACT

Environmental isotopes can be used to test whether integrated water quality models represent catchment processes appropriately. We review several case studies and describe how cesium-137 and beryllium-7 can be used to identify which erosion and deposition processes operate in the landscape (eg. interrill, rill, ephemeral gully erosion), and how oxygen-18 and deuterium can be used to determine the relative importance of the various storm runoff generation mechanisms in catchments.

INTRODUCTION

With increasing frequency, land and water resource managers are asked to evaluate existing water resource degradation problems, to predict the consequences of proposed remedial measures, and to forecast the effects of changes in resource use. These managers base their decisions largely on "integrated models", models integrating the various processes in the hydrologic cycle and the transport of agricultural chemicals and sediment (DeCoursey 1985). In this review, we will focus on two aspects of integrated models for agricultural nonpoint-source problems: (1) erosion and deposition processes and (2) runoff generation processes; both in the context of storm runoff events.

During the past few decades, hydrologists have been learning to measure and interpret differences in the concentrations of environmental isotopes to assess the temporal and spatial variations in flow and transport processes. Environmental isotopes are naturally occurring isotopes or isotopes already introduced into the environment by man, which can be used to study hydrologic processes. Our main objective in this report is to describe how environmental isotopes can be used to identify which hydrologic processes are important in problems concerned with nonpoint-source pollution in agricultural catchments. Specifically, we will discuss the use of cesium-137 and beryllium-7 in erosion and deposition studies, and oxygen-18, deuterium, and tritium in runoff generation studies.

Finally, we will suggest how isotopic techniques can test whether a particular model accurately represents the combination of catchment processes.

[1]M.G. Sklash, Associate Professor, Department of Geology, University of Windsor, Windsor, Ontario, Canada N9B 3P4.

[2]I.D. Moore, Associate Professor, Department of Agricultural Engineering, University of Minnesota, St. Paul, MN 55108.

[3]G.J. Burch, Assistant Director, Bureau of Rural Science, Department of Primary Industries and Energy, Barton, A.C.T. 2600, Australia.

MODEL ATTRIBUTES REQUIRED BY USERS

Integrated models are useful to both resource managers and researchers. Resource managers include land and water users and government planners whose mandate is to ensure that specificresources are effectively managed, to maintain their integrity, and to sustain their usefulness or productivity. Land and water resources must be managed at a variety of scales progressing from field, to catchment, to regional levels. Integrated models must either be adaptable to each of these scales or separate models must be selected to suit the particular scale under consideration.

Relatively few models are available that encompass both the full scope of surface and groundwater hydrology for whole catchments and the associated processes of sediment and solute transport. The models that have been derived deal with the component catchment processes in one of two ways. "Lumped parameter" models represent discrete physical processes empirically and then extrapolate areally using functional relationships which do not simulate specific interacting physical processes. "Physically-based process" models attempt to simulate the real time and scale distribution of physical processes taking place in the catchment (DeCoursey 1985).

Resource managers usually have little concern for the integral components of models provided that the effects of the management options they wish to evaluate can be accurately simulated by the models. In contrast, researchers are far more concerned with the selection of appropriate simulation procedures and process descriptions that give detailed insights into the behavior of single or interacting processes or of entire systems. The lumped-parameter models are more likely to suit the resource manager, while physically-based models are more likely to be used by researchers. The need for these models and the procedure for selecting appropriate mathematical models for application to nonpoint-source water pollution control are described by DeCoursey (1985).

STREAMFLOW AND RUNOFF GENERATING MECHANISMS

To successfully integrate components of models that can simulate water movement in catchments as well as the transport of particulates and solutes, we must understand the mechanisms of streamflow and runoff generation. Review papers by Freeze (1974); Dunne (1978, 1983); Ward (1984) and others attempt to sort out the various processes. Some mechanisms are more important in certain landscapes than others depending on soil, geology, vegetation and topographic characteristics. These mechanisms can be complex and despite a high level of research input, researchers disagree about the roles of certain mechanisms (Pearce et al. 1986). In reviewing these mechanisms, we will make no attempt to resolve the differing opinions. Nevertheless, to simulate catchment behavior realistically, it is understood that models must incorporate functional relationships which correctly employ established physical principles and known hydrological processes.

The conclusion of Freeze (1974) has been consistently supported by whole-catchment investigations. That is, three principal mechanisms generate storm runoff from major precipitation events:

(1) overland flow on areas saturated from above by infiltrating rain (Hortonian overland flow),
(2) overland flow on areas saturated from below by a rising water table (saturation overland flow), and
(3) subsurface flow.

The generation of stream quick flow (Hewlett and Hibbert 1967) after major rainfalls, accordingly, can involve both surface and subsurface components Dunne (1983).

Runoff generation is not simple. Burch et al. (1987) have demonstrated that before reaching a stream channel, water may follow one of a number of alternative flow paths or several paths sequentially. Whole-catchment studies covering a wide variety of environmental, geomorphological, soil, and vegetative conditions indicate that the balance of contributions to storm flow from the above three mechanisms can vary widely (Moore et al. 1986, Pearce et al. 1986, Burch et al. 1987). Therefore, to simulate storm responses realistically, integrated models must accommodate the complex interaction of the above catchment characteristics to reproduce experimentally-observed responses. Already models have emerged, such as ANSWERS (Beasley et al. 1980), TAPES (Moore et al. 1988b), and TOPMODEL (Beven and Kirkby 1979, Beven and Wood 1983) that show preliminary promise in predicting catchment storm responses when overland flow mechanisms have a dominant influence.

ISOTOPIC TRACERS FOR HYDROLOGIC AND EROSION RESEARCH

The principle attributes of a useful tracer for detecting water, solute, or sediment movement on a catchment scale are that the tracer:

(1) be readily measured using standard laboratory procedures;
(2) be chemically and physically stable and not subject to environmental degradation or transformation;
(3) have inputs to the environment that can be quantified;
(4) be uniformly distributed during the input phase and show no bias during transport; and
(5) occur as a natural component of, or display a specific affinity for, the media under investigation.

Most tracers in hydrologic and erosion studies may be classified as either chemical or isotopic. Isotopic tracers generally satisfy more of the above attributes than do chemical tracers and environmental isotopes fulfil even more of the required attributes. In this paper we will present case studies describing some of the most useful environmental isotope tracers for studying sediment transport and storm runoff generation.

Cesium-137 and Beryllium-7 in Erosion Research

Empirical models for erosion prediction, such as the USLE model (Wischmeier and Smith 1978), incorporate the range of factors (vegetative cover, slope, slope length, soil surface condition, erodibility and rainfall amount and intensity) affecting runoff, sediment entrainment, and transport on uniform slopes. In the more recent mechanistic models, these processes are represented by physically-based mathematical expressions which also take into account infiltration, runoff, overland flow, sediment entrainment, transport and deposition (Rose et al. 1983b, Foster et al. 1987). The wealth of information now available on erosion and deposition processes is currently enabling modelers to more confidently apply computer simulation to predict the sediment dynamics of reasonably complex landscapes (Beasley et al. 1980). However, testing the predicted movement of sediment within a catchment or landscape with spatially heterogeneous vegetative and soil attributes is difficult or prohibitive using traditional measurement techniques. Isotopic tracers offer a means of model testing and further investigation of sediment movement processes. For these purposes, two isotopes are very useful: cesium-137 (^{137}Cs), which is well-researched, and another less well-known isotope, beryllium-7 (^{7}Be).

^{137}Cs originated from the release of fission products into the atmosphere, with most production related to atmospheric testing of nuclear weapons in the 1950's and 1960's. ^{137}Cs is characterized by a relatively short half-life (30.2 years) and strong cationic adsorption to both mineral and organic components of soil (Lomenick and Tamura 1965). Rainfall infiltration causes surface accumulations of ^{137}Cs in undisturbed soils. ^{137}Cs profiles are characterized by concentrations which decrease with depth, becoming undetectable within centimeters to tens of centimeters below the surface (Ritchie et al. 1974). Eroded sites can be identified by the loss of this surface accumulation of the isotope, whereas deposition sites have buried or substantially deeper isotopic profiles. Since current rates of ^{137}Cs deposition are relatively low (consisting mostly of residual ^{137}Cs, originally produced during widespread atmospheric testing of nuclear weapons which terminated in 1963), sites under investigation must have an inventory of the isotope to allow soil loss or gain to be detected. Some highly eroded sites may now have insufficient ^{137}Cs remaining in the soil to obtain useful measurements.

Burch et al. (1988) have improved the isotopic technique of erosion and deposition process research by interpreting measurements of multiple tracer isotopes in soils and sediments. They demonstrate that ^{7}Be, with its different half-life and different depth profile in soil and sediment, is particularly useful as a complementary isotope in ^{137}Cs studies.

^{7}Be, naturally produced by cosmic ray spallation of nitrogen and oxygen in the upper atmosphere, has a half-life of 53.3 days (Young and Silker 1980). Since ^{7}Be production is natural, deposition at a site is relatively constant. Rainfall timing and amount are important: higher ^{7}Be activities occur in rain after dry periods and during brief storms than in long-duration rainfalls. Since ^{7}Be is strongly adsorbed to the mineral and organic soils, detectable ^{7}Be is confined to the upper 5 mm or so of the soil profile.

Since both isotopes strongly adsorb to soil particles, they concentrate in the surface layers. Profile measurements, therefore, will indicate where erosion or deposition has occurred. As a result of its longer half-life, ^{137}Cs penetrates the soil to greater depths than does the short-lived ^{7}Be. ^{137}Cs and ^{7}Be concentrations in a soil or sediment can be jointly examined to assess the nature of erosion and deposition patterns in a catchment. For example:

(1) low levels or absence of both isotopes in soil indicate erosion has occurred;
(2) a sediment sample which is high in both ^{137}Cs and ^{7}Be must have originated by "sheet" erosion from the upper few millimeters of soil;
(3) a sediment derived from "rilling" will have high ^{137}Cs activity but low ^{7}Be activity;
(4) a sediment high in ^{7}Be but low in ^{137}Cs is from a site that was deeply eroded but subsequently exposed long enough to develop an inventory of ^{7}Be, for example, a gully; and
(5) a sediment low in both ^{137}Cs and ^{7}Be may originate from the collapse of a gully wall.

Burch et al. (1988) report on an erosion study in which ^{7}Be and ^{137}Cs were measured in soils and sediments from two small agricultural watersheds at Wagga Wagga, New South Wales, Australia. Their "untreated" 7.3-ha catchment has been uncultivated for over 40 years, and is gullied; the "treated" 7.5-ha catchment was treated 40 years ago to reduce erosion, and recently was cultivated in wheat. Observations were also made on soils and sediments from a number of adjacent 22m-long runoff/erosion plots given various tillage and cropping treatments. The mean annual precipitation is about 580 mm and the average annual runoff is 38 mm from the untreated catchment and 12 mm from the treated catchment. The average annual soil loss, based on a 40-yr record, is highly variable from the untreated catchment with an average loss of 2.0 t/ha, compared to an average 0.03 t/ha from the treated catchment.

The catchments and plots were equipped to measure rainfall, discharge, and sediment. Small sediment traps were installed below the H-flume outlets to collect most, and sometimes all, of the sediment leaving the catchments. Runoff and sediments were also collected for each storm event from the small experimental plots using discharge splitters and collector tanks. The collection of sediment samples from sequential storm events throughout 1986 represent a wide range of soil sources including: gully walls, well-grassed pasture soils, stable cultivated soil with complete crop cover, and rill eroded cultivated soil.

^{7}Be activities at the soil surface ranged from as low as 0.1 Bq/kg for eroded fallow soils to 4.4 Bq/kg for noneroded pasture soils. The high ^{7}Be activities of the sediments (fig. 1) indicate that small amounts of erosion of generally stable surface soils by sheet flow can yield sediments very highly enriched with this isotope. The sediments derived from actively eroding surfaces have a wide range of ^{7}Be activities, and most sediments are highly enriched relative to the source soils. This enrichment indicates that active scavenging for ^{7}Be must occur during sediment entrainment and transport, even in the short (22-m) runoff/erosion plots. However, there were also several occurrences of distinctly low ^{7}Be activities in sediments that correspond with major sediment inputs caused by the collapse of gully walls or severe wash from rill eroded fallow plots.

The ^{137}Cs profiles in the soils were as expected with progressively lower concentrations with depth and with few measurable concentrations deeper than 50 mm. ^{137}Cs activities in the sediments showed no enrichment relative to the source soils. These trends reflect the current very low levels of atmospheric input, the extended period of deposition, and a deeper distribution through the upper 50-100 mm of soil (relative to ^{7}Be).

The variation in ^{7}Be and ^{137}Cs activities in the sediments is indicative of the hydraulic behavior at the eroding surface, that is sheet flow, rill flow, or gully flow. Sediments with low ^{7}Be and low to moderate ^{137}Cs concentrations were derived from gully wall sites in an eroded catchment and from deeply incised rills in fallow plots, after major runoff and erosion events. Sediments with low to moderate ^{137}Cs levels, but moderately enriched in ^{7}Be, were derived from catchment gully floor

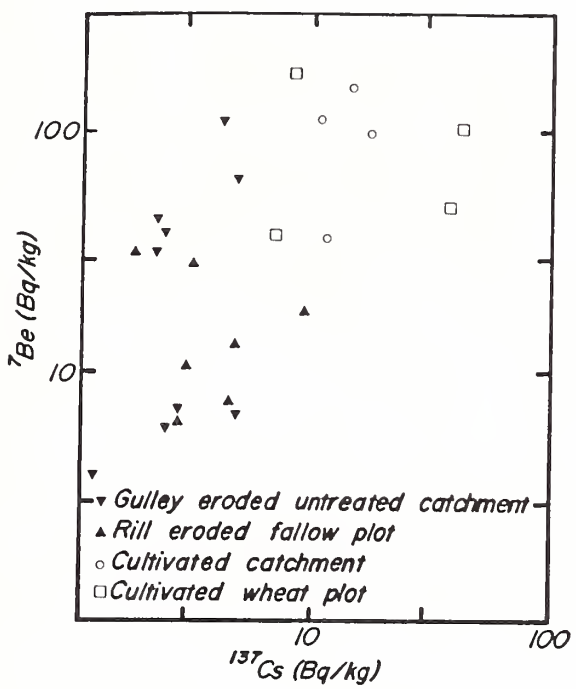


Figure 1.
Measured ^{7}Be and ^{137}Cs activities in sediments from Wagga Wagga, New South Wales, Australia (after Burch et al. 1988).

deposition areas and rill-eroded fallow plots, after small to intermediate storm events.

Sediments high in both ^{7}Be and ^{137}Cs were consistently derived from soils protected by pasture or crop vegetation and subjected to sheet erosion over a range of storm intensities. Therefore, isotopic activities in sediments and soils provide a means of quantitatively tracing the origin of the sediments and testing the prediction of erosion-deposition models.

Oxygen-18, Deuterium, and Tritium in Runoff Generation Research

During the past two decades, oxygen-18 (^{18}O), deuterium (D), and tritium (T) have been widely used to trace water movement through catchments. Two main attributes make these isotopes particularly effective tracers: (1) they are constituent parts of natural water molecules and therefore travel at the same rate and along the same flow pathways as "average" water, and (2) they are chemically conservative at the low temperatures associated with most small watershed systems (Fritz et al. 1976, Kennedy et al. 1986, and others), i.e. their concentrations in a volume of water do not change by reactions with the catchment materials. The conservative nature of these isotopes makes them superior to natural chemical tracers which could change in concentration by reactions with catchment materials.

This paper only briefly reviews the environmental isotope hydrograph separation technique. Details on the distributions of ^{18}O, D, and T in the hydrosphere are given in Payne and Halevy (1968), Fritz and Fontes (1980), Gat and Gonfiantini (1981), Faure (1986), and Hoefs (1987). Rodhe (1987) and Pearce et al. (1986) give extensive discussions on the hydrograph separation technique.

Natural water contains about 2,000 ppm of molecules with ^{18}O and 320 ppm of molecules with D (Payne and Halevy 1968). Variations in concentrations of ^{18}O and D in precipitation stem largely from temperature-dependent isotopic fractionation during condensation, and different origins of air masses, however, ^{18}O and D concentrations are linearly related. The concentrations of ^{18}O and D, determined using mass spectrometers, are quoted as "delta" or "del" values (δ) which are per mil ($^{0}/_{00}$) differences relative to a standard according to:

$$\delta = [(R_x - R_{SMOW})/R_{SMOW}] \times 1000\ ^{0}/_{00} \qquad [1]$$

where δ = $\delta^{18}O$ or δD
R = $^{18}O/^{16}O$ or R = $D/^{1}H$,
x = unknown water sample, and
SMOW = Standard Mean Ocean Water (Craig 1961).

Tritium is a radiogenic isotope of hydrogen with a half-life on the order of 12.4 years (Fritz and Fontes 1980). Concentrations of T in precipitation are expressed in "tritium units" (TU) where one TU is equivalent to one T atom in 10^{18} hydrogen atoms. Cosmogenically-produced T concentrations in precipitation are estimated to be on the order of 4-25 TU (few measurements had been made prior to the massive anthropogenic inputs). The advent of large-scale atmospheric testing of nuclear devices in 1952 increased T concentrations in precipitation by several orders of magnitude. Tritium concentrations in precipitation peaked in 1963 at several thousand TU in the northern hemisphere. Since the atmospheric test ban treaty was signed, concentrations have gradually declined (Gat 1980).

^{18}O, D, and T are used to separate storm (and snowmelt) hydrographs based on the distinctive isotopic signatures carried by the contributing components. Distinctive isotopic signatures develop as water molecules pass through the catchment via different flow paths with different residence times.

The most widely used isotopic hydrograph separation technique employs a two-component mixing model. The model assumes that water in a stream during a runoff event is a mixture of two components: "new" water, from the current rain or snowmelt event, and "old" water, subsurface water already present in the catchment prior to the current rain or snowmelt event. The isotopic signature of the "old" water develops from a mixing of precipitation which infiltrates over weeks, months, or years. The "new" water isotopic signature can vary seasonally, between storms, and during storms for ^{18}O, D, and T; and in the case of T (due to the atmospheric test ban treaty and its half-life), there has also been the gradual decline since 1963. Figure 2 demonstrates the potential difference in $\delta^{18}O$ values between "old" water (shallow groundwater and stream baseflow) and "new" water (the seasonal variation in precipitation) in southwestern Ontario, Canada. It follows from figure 2 that some storms will have the same $\delta^{18}O$ value as the "old" water and that these storms could not be separated using ^{18}O (unsuitable storms could also occur for D and T).

During baseflow periods, all stream water is discharged groundwater, and its isotopic signature is an integration of upstream "old-water" discharges. During storm and snowmelt runoff events, however, "new water" is added to the stream. If the "old" and "new" water components are isotopically different and the "new" water component remains isotopically constant during the event, the stream water becomes diluted by the addition of the "new" water, with the extent of dilution dependent on the relative contributions of the "old" and "new" water components. The contributions of "old" and "new" water are calculated by solving mass balance equations provided that the stream, "old water", and "new water" tracer concentrations are known (fig. 3). These equations are expressed as:

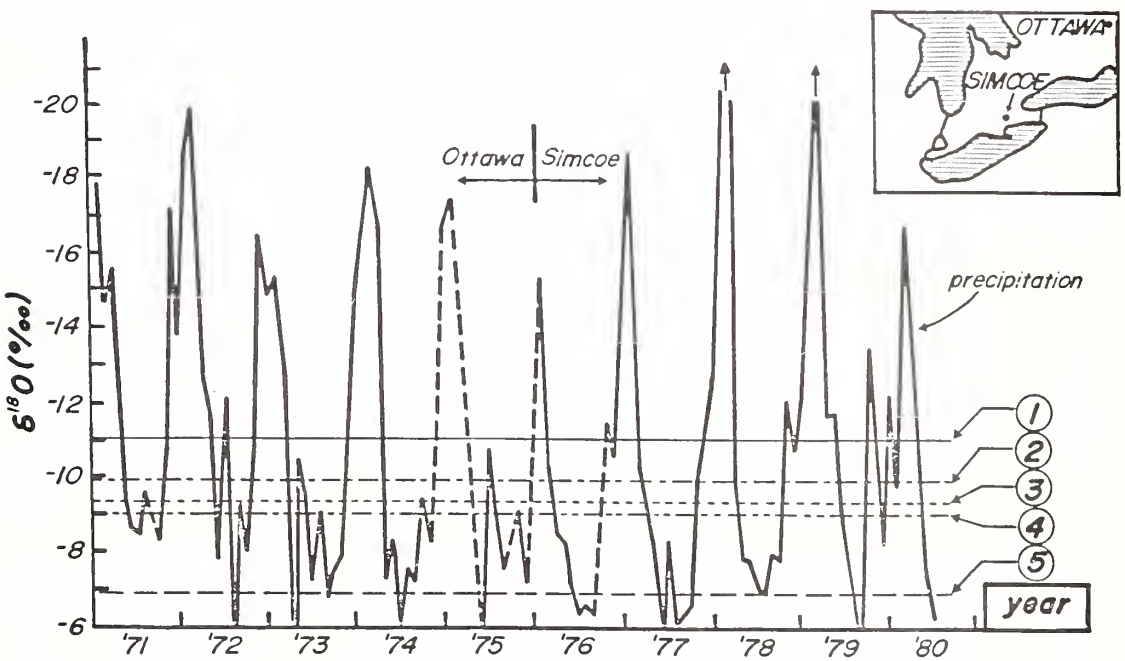


Figure 2.
^{18}O concentrations in precipitation, shallow groundwater and stream baseflow in southwestern Ontario, Canada. Precipitation data from International Atomic Energy Agency reports and Fritz and Drimmie (personal communication), baseflow data from Sklash (1978), Fritz et al. (1976), and Sklash (unpublished data): 1 = baseflow in Canagagigue Creek in February 1977; 2 = baseflow in Big Otter Creek in May 1974; 3 = shallow groundwater in Sarnia in May 1986; 4 = baseflow in Sturgeon Creek in February 1985; 5 = baseflow in Hillman Creek in April 1977.

$$Q_s = Q_o + Q_n \quad [2]$$

$$C_sQ_s = C_oQ_o + C_nQ_n \quad [3]$$

which can be solved for Q_o:

$$Q_o = [(C_s - C_n)/(C_o - C_n)]\, Q_s \quad [4]$$

where Q = discharge,
C = the isotopic concentration of a component or the stream, and
s, n, and o = subscripts indicating the "stream", "new" water component, and the "old" water component, respectively.

Although the two-component mixing model for stream hydrograph separation calculates the proportions of "old" and "new" water in the stream, the technique does not identify the actual flow path(s) of the water. It does, however, allow comparative evaluation of the runoff generation processes in a catchment. For example, if overland flow or subsurface flow through macropores dominates runoff generation, an isotopic study should detect mostly "new" water in the stream. Conversely, if Darcian subsurface flow dominates storm runoff production, the isotopic study should detect mostly "old" water in the stream. Evaluation of the flow processes can be enhanced by applying the mixing model to water collected from subsurface runoff collection pits, overland flow, and macropores.

The hydrograph separations are also useful for estimating the areal extent of overland flow contributing areas (Rodhe 1987, Pearce et al. 1986, and Fritz et al. 1976) and the mean residence

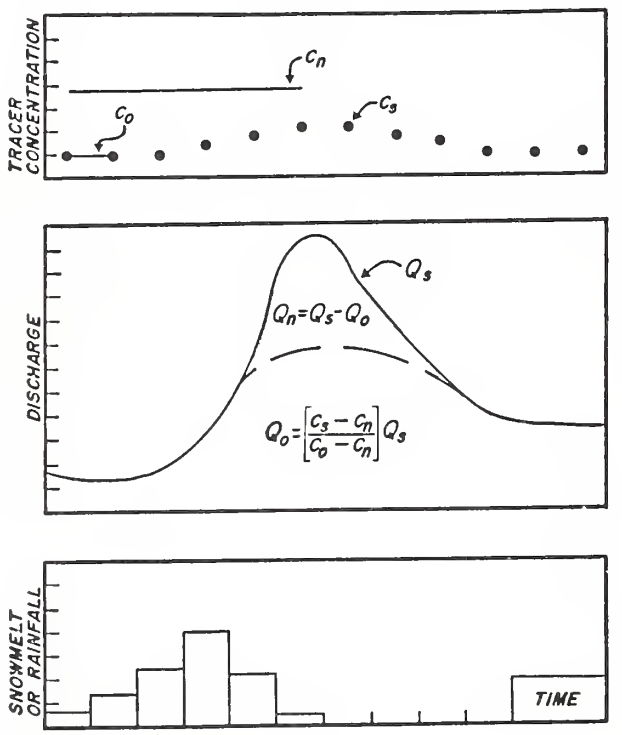


Figure 3.
Hydrograph separation using environmental isotope tracers - a hypothetical example.

time of the "old" water component (Maloszweski et al. 1983, Pearce et al. 1986, Turner et al. 1987, and others). The area generating overland flow is estimated by dividing the volume of "new" water in storm runoff by the net precipitation (i.e., gross precipitation minus interception) for the event. The mean residence time of the "old" water in a catchment can be estimated by comparing the seasonal variations in del values of the input (precipitation) and the output (streamflow) in a catchment or by comparing the long-term T input and output functions.

Hydrograph separation studies using environmental isotope techniques have been reported for more than 30 catchments. By far the most important conclusion is that "old" water usually dominates storm and snowmelt runoff events in streams. This conclusion implies a significant and rapid displacement of "old" water to the stream. The physical mechanism is discussed in Sklash et al. (1986) but is beyond the scope of this paper. Here, we will briefly review two useful case histories: the Hillman Creek study (Sklash and Farvolden 1979), which was part of a study on nitrate migration in agricultural watersheds, and the Maimai study (Pearce et al. 1986 and Sklash et al. 1986) which clearly demonstrates the value of using isotopes in determining runoff generation mechanisms.

The first case study, the Hillman Creek catchment, consists of 100 ha of agricultural land in southwestern Ontario, Canada. The catchment has low relief (<10 m), 1-5 m of sandy soil over glacial till, a water table depth of 1-4 m, intensive agriculture, few trees, and exposed soil. Annual precipitation averages 740 mm and is evenly distributed over the year. Sklash and Farvolden (1979) monitored several major runoff events in Hillman Creek during 1977, two of which are briefly discussed below.

Rainfalls of 38, 35, and 31 mm fell on the catchment on April 22, 23, and 25, respectively, causing stream discharge to increase from about 6 L/s on April 22 to at least 85 L/s on April 23 and 90 L/s on April 25. Even though baseflow $\delta^{18}O$ prior to the storms was about $-6.9^{0}/_{00}$ and the rain on April 25 had a weighted averaged $\delta^{18}O$ of $-18.1^{0}/_{00}$, the stream $\delta^{18}O$ value never became more negative than $-9.0^{0}/_{00}$. These observations indicate that "old" water provided about 80% of the storm runoff volume. This magnitude of "old" water contribution is consistent with "old" water estimates based on groundwater stage-groundwater discharge rating techniques, hydrograph separation based on electrical conductivity values, and isotopic and electrical conductivity observations of overland flow. Most of the storm runoff events on Hillman Creek were dominated by "old" water; the only exception occurred on June 6.

On June 6, following a very dry period, an intense thunderstorm dropped 36 mm of rain on the catchment within 2 hours. Stream discharge quickly rose from 5 L/s to over 200 L/s. All of the data from this event indicated that "new" water dominated storm runoff: the $\delta^{18}O$ and electrical conductivity values of both the stream and overland flow approached the "new" water value and the groundwater stage-groundwater discharge data indicated that the stream temporarily became influent.

Observed nitrate concentrations in the stream during these two events were consistent with the runoff components indicated by the isotopic data (Sklash 1978). During the April 25th event, which was dominated by "old" water, the stream nitrate concentrations rose considerably above baseflow concentrations. The stream nitrate increase was consistent with observations of abundant high nitrate overland flow which originated as groundwater discharge from remote seeps. Environmental isotopes were instrumental in determining that the overland flow was dominantly "old" water during this event. In contrast, during the June 6th event, when both the stream and overland flow were dominated by "new" water, the stream nitrate concentrations decreased from the baseflow level until after peak stream discharge. During this event, the observed overland flow had low nitrate concentrations and $\delta^{18}O$ values characteristic of "new" water.

The second case study involves the Maimai Experimental Catchments that are hydrologically responsive, forested catchments on the South Island of New Zealand (Pearce et al. 1986 and Sklash et al. 1986). These small (<4 ha) catchments are characterized by short (<150 m), steep slopes (average 34°) with thin (<1 m), permeable soils. Both Pearce and McKerchar (1979) and Mosley (1979) agree that only 4-7% of the catchment area is capable of producing saturation overland flow and that storm runoff is dominated by subsurface flow. There has been considerable debate, however, about the nature of the subsurface flow: is it "old" or "new" water?

On the basis of dye tracer tests and analysis of discharge hydrographs from throughflow collection pits, Mosley (1979, 1982) concluded that subsurface flow during storms is dominated by macropore flow of "new" water. Pearce et al. (1986) and Sklash et al. (1986) analyzed rain, streamwater during baseflow and storm runoff periods, groundwater, and samples of throughflow for ^{18}O, D, chloride, electrical conductivity, and pH. Their studies clearly demonstrate that "old" water dominates storm runoff in the Maimai catchments. More generally, their results indicate that hydrometric studies and chemical tracer studies alone can lead to incorrect conclusions about storm runoff generation mechanisms and residence times.

Data from isotopic hydrograph separation studies have only recently been incorporated into modeling efforts (Christopherson et al. 1985, Bobba et al. 1985). Most of the work so far is related to integrated models to deal with the fate of acid precipitation in acid-sensitive catchments.

ISOTOPIC TESTING OF MODELS

If an integrated water quality model accurately portrays catchment processes (i.e. sheet, rill or gully erosion; and Hortonian overland flow, saturation overland flow, or subsurface flow), it should be possible to use the measured temporal variations in stream isotopic character during an event (^{18}O, D, and T) and the isotopic character of eroded areas and accumulated sediments (^{137}Cs and ^{7}Be) after an event, to test the predictions of the model and whether the magnitude of the various processes predicted by the model are correct.

Tables 1 and 2 list several "bulk water transport" and "erosion and sediment yield" models that could be tested using environmental isotope tracer techniques. Although the tables do not list all models reported in the literature, they contain examples of the various categories of models. Table 1 lists bulk water transport models that in some way represent the effects of saturated source areas on runoff generation. Table 2 examines erosion and sediment yield models that deal dominantly with upland erosion, i.e. the table does not include models which are largely oriented to erosion and sediment transport in streams. In both tables 1 and 2, we have highlighted the modeled catchment processes which could be tested using environmental isotopes.

Once an isotopic study has been conducted in a particular catchment, two questions can be asked about the model to be used for that catchment. Does the model include all the processes which are identified by the isotopic study? Does the model determine the proportions of each process indicated by the isotopic study? In other words, are the modeled processes consistent with the isotopic data? We can use two models from tables 1 and 2 to demonstrate.
O'Loughlin's (1981) quasi-two-dimensional flow model attributes storm runoff to some combination of saturation overland flow and saturated subsurface flow. From an isotopic viewpoint, saturated subsurface flow contributions, by definition, are dominated by "old" water; and although the contributions of "old" and "new" water in saturation overland flow have not been tested, saturation overland flow implies "new" water dominance. O'Loughlin's model prescribes particular proportions of the two components and predicts a storm runoff hydrograph based on the characteristics of the watershed and of the particular runoff-producing event. It should be possible to predict the temporal variations in stream isotopic character during each event by rearranging equation 3 and solving for C_s, the isotopic character of the stream:

Table 1.
Runoff generation processes included in selected bulk water transport models

MODEL CLASSIFICATION, AUTHOR, AND CODE NAME	TYPE OF SOLUTION[1]	RUNOFF PROCESSES[2] HOF	SOF	SSF	USF
ONE-DIMENSIONAL HILLSLOPE					
Henderson & Wooding (1964), Childs (1971), Chapman (1980), O'Loughlin (1981), Sloan & Moore (1983)	A		X	X	
Beven (1981), Nieber (1982)	N-fe		X	X	X
Beven (1982)	A	X	X	X	X
Nieber (1982)	N-fe	X	X	X	X
Sissin et al. (1980), Hurley & Pantelis (1985), Stagnitti et al. (1986)	A		X	X	X
TWO-DIMENSIONAL HILLSLOPE					
Neuman (1973) UNSAT2, Reeves & Duguid (1975), Beven (1977), Mohsenisaravi (1981), Nieber & Walter (1981), Davis & Neuman (1983)	N-fe	X	X	X	X
Freeze (1972b), Stephenson & Freeze (1974)	N-fd	X	X	X	X
Troendle (1972)	N-fd		X	X	X
Smith & Woolhiser (1971)	N-fd	X			
O'Loughlin (1981)	A		X	X	
THREE-DIMENSIONAL HILLSLOPE					
Freeze (1971, 1972a, 1978)	N-fd	X	X	X	X
Smith & Hebbert (1983)	N-fd	X	X	X	X
LUMPED- AND DISTRIBUTED-PARAMETER CATCHMENT					
Crawford & Linsley (1966) SWM	D	X		X[3]	X[3]
Federer & Lash (1978) BROOK, Moore et al. (1983)	L	X	X	X[3]	X[3]
Beasley & Huggins (1982) ANSWERS	D	X		X[3]	X[3]
Thomas & Beasley (1986a, 1986b) ANSWERS	D	X	X	X[3]	X[3]
Abbott et al. (1986a,b), Bathurst (1986a,b) SHE	D	X	X	X	X
DIGITAL TERRAIN/RELIEF-BASED MODELS					
Beven & Kirkby (1977), O'Loughlin (1986), Moore et al. (1988b)	DT		X		
Beven & Kirkby (1979), Beven & Wood (1983) TOPMODEL, Moore et al. (1986b)	DT		X	X	

[1]A = analytical solution, N = numerical solution (fe = finite element, fd = finite difference), L = lumped-parameter, D = distributed-parameter, DT = digital terrain model

[2]HOF = Hortonian overland flow, SOF = saturation overland flow, SSF = saturated subsurface flow, USF = unsaturated subsurface flow

[3]no real differentiation between saturated and unsaturated flow

Table 2.
Erosion processes included in selected erosion and sediment yield models.

MODEL CLASSIFICATION, AUTHOR, AND CODE NAME	TIME SCALE[1]	PROCESSES[2] Detach	Trans	Depos	TYPE OF EROSION Sheet	Rill	Gully
FIELD-SCALE MODELS							
Wischmeier & Smith (1978) USLE	A	X			X[3]	X[3]	
Williams (1975b, 1978) MUSLE	E	X			X[3]	X[3]	
Knisel (1980) CREAMS	E[4]	X	X	X	X	X	
Rose et al. (1983b,c) GUESS	E	X	X	X	X[3]	X[3]	
Foster et al. (1987), Lane et al. (1987), Gilley et al. (1987) WEPP	E	X	X	X	X	X	
CATCHMENT-SCALE MODELS							
Wilson et al.(1984a,b) SEDIMOT-II	E	X	X	X	X	X	
Donigian et al. (1977), Donigian & Davis (1978) ARM, Donigian & Crawford (1976) NPS, Leytham & Johanson (1979) WEST	E, C	X	X	X	X	X	X
Beasley et al. (1980), Beasley & Huggins (1982), Storm (1986), Dillaha & Beasley (1983) ANSWERS	E	X	X	X	X	X	
Young et al. (1987, 1989) AGNPS	E	X	X	X	X[3]	X[3]	
Ahnert (1976, 1987) SLOP 3D	G	model of land form development					
DIGITAL TERRAIN / RELIEF-BASED MODELS							
Moore & Burch (1986)	S, DT		X	X	X		
Thorne et al. (1986), Moore et al. (1988a)	DT						X[5]

[1]Detach(ment), Trans(port), Depos(ition)

[2]A = average annual, E = storm event, C = continuous, G = geologic time, S = steady state, DT = digital terrain

[3]indicates types of erosion are not differentiated

[4]hydrology component of CREAMS is continuous

[5]ephemeral gully erosion

$$C_s = (C_oQ_o + C_nQ_n)/Q_s \quad [5]$$

where Q_o and Q_n now represent contributions of saturated subsurface flow ("old" water) and saturation overland flow ("new" water), respectively, predicted by O'Loughlin's model, and the other parameters are as defined previously. O'Loughlin's model would be considered to properly represent the physics of the system if it consistently predicted C_s values which match observed stream isotopic concentrations.

ANSWERS (Beasley et al. 1980 and others) is an event-based erosion and sediment yield model which categorizes erosion as either sheet (interrill) or rill erosion. The efficacy of this model could be tested using ^{7}Be and ^{137}Cs as tracers. For example, if the model is correct in predicting that only sheet erosion occurs during a given storm on a given catchment, not only should the observed and predicted amounts of sediment match, the eroded area should be stripped of ^{7}Be and a thin surface layer of ^{137}Cs. Also, the sediment transported to and measured at the catchment outlet should have high levels of both ^{7}Be and ^{137}Cs. However, if the model prediction is incorrect and rill erosion were the dominant mechanism, sediments would have high ^{137}Cs concentrations and low to moderate ^{7}Be concentrations. This would be an example of a model that can handle all of the appropriate processes but with an incorrect distribution of processes.

If an erosion event yields sediments that are derived from gully wall erosion and collapse, the sediments would have both low ^{137}Cs and ^{7}Be activities. In this case, ANSWERS might correctly compute the sediment and water fluxes at the catchment outlet, but because ANSWERS does not account for gully processes, it would incorrectly represent the relative magnitude of the various erosional processes in that landscape. This would be an example of a model that cannot calculate the correct distribution of processes because it does not include all of the processes. In particular, the model would incorrectly predict the distribution of sediment sources across the catchment. Because of this, the model might not be suitable for micro-targeting soil conservation remedial measures in the catchment, even though the predicted sediment and water fluxes at the catchment outlet were correct.

Wall et al. (1979) and Dickinson and Wall (1977) have provided evidence linking the source area concept of runoff generation to the erosional behavior of landscapes. In a study of two small agricultural catchments in Canada, Wall et al. (1979) observed that (i) sediment source areas constitute a relatively small proportion of the total catchment, (ii) these areas are temporally highly variable, and (iii) they are in close proximity to streams. Hence, runoff generation and erosion processes active in the landscape are highly interdependent. To correctly predict the magnitudes and relative contributions of the various erosion processes, it is essential that the magnitudes and pathways of water transport also be accurately predicted. As described above, the analysis of a range of environmental isotope tracers has the capability of verifying the interactions and processes.

CONCLUDING REMARKS

Over the past two decades, hydrologists have increasingly turned to environmental isotopes to characterize catchment processes. In this review, we have demonstrated the usefulness of combining ^{137}Cs and ^{7}Be surveys in locating areas of erosion and deposition and in determining the type of erosion which has occurred. We have also reviewed two case studies which show how the concentrations of ^{18}O and D in water can be interpreted to evaluate mechanisms of storm runoff generation in a watershed. In these case studies and in other studies like them, environmental isotopes have provided information on runoff generation and nonpoint-source pollutants which would be difficult to attain even with very intensive instrumentation. We

recognize that environmental isotopes have a great potential for unraveling the complex spatial and temporal variations in catchment behavior.

Even though integrated water quality models have merit in predicting flow and/or transport, they are often most effective for simulating laboratory, plot, and small-catchment scales. To optimize model performance on catchment and regional scales, it is our contention that the catchment processes addressed in models should be consistent with the catchment processes identified using isotopic techniques.

In summary, environmental isotope measurements should be considered in hydrologic modeling: (1) to give a clearer picture of how catchments operate to ensure that the appropriate catchment process(es) are incorporated into the structure of new models, (2) to test if an existing model applied to a given catchment includes all of the appropriate catchment processes, and (3) to test if an existing model which includes all of the appropriate catchment processes, uses the appropriate distribution of processes.

ACKNOWLEDGMENTS

This paper is published as Paper No. 16,366 of the Scientific Journal Series of the Minnesota Agricultural Experiment Station on research conducted under Minnesota Agricultural Experiment Station Project No. MIN-12-55. The hydrograph separation work by Sklash has been funded by grants from the Natural Sciences and Engineering Research Council of Canada and supported by the Forest Research Institute (New Zealand) and the Institute of Nuclear Sciences (New Zealand). Dr. A.J. Pearce and Dr. M.K. Stewart played important roles in the New Zealand work. The work by Moore was supported in part by a grant from the Australia National Soil Conservation Program and a grant from the Graduate School of the University of Minnesota. The research by Burch has been financially supported by Australia National Soil Conservation Program. Isotopic analyses were provided by the Research School of Earth Science, Australian National University and the Department of Applied Science, Royal Melbourne Institute of Technology. The authors thank these supporting organizations and Dr. G.R. Foster and Dr. J.L. Nieber for reviewing the manuscript and providing many helpful suggestions. The reviewers, particularly Doug James from Utah State University, supplied us with many useful comments.

LITERATURE CITED

Abbott, M.B., J.C. Bathurst, J.A. Cunge, and others. 1986a. An introduction to the European Hydrological System - Systeme Hydrologique Europeen, "SHE", 1: history and philosophy of a physically-based, distributed modelling system. J. Hydrol. 87:45-59.

Abbott, M.B., J.C. Bathurst, J.A. Cunge, and others. 1986b. An introduction to the European Hydrological System - Systeme Hydrologique Europeen, "SHE", 2: structure of a physically-based, distributed modelling system. J. Hydrol. 87:61-77.

Ahnert, F. 1976. Brief description of a comprehensive three-dimensional model of a landform development. Z. Geomorph. N.F., Suppl.-Bd. 25:29-49.

Ahnert, F. 1987. Uses of a theoretical model in the analysis of landform development. In A. Godard, and A. Rapp, eds., Processes and Measurement of Erosion, pp. 555-559. Editions du Centre National de la Recherche Scientifique, Paris.

Bathurst, J.C. 1986a. Physically-based distributed modelling of an upland catchment using the System Hydrologique Europeen. J. Hydrol. 87:79-102.

Bathurst, J.C. 1986b. Sensitivity analysis of the Systeme Hydrologique Europeen for an upland catchment. J. Hydrol. 87:103-123.

Beasley, D.B., and L.F. Huggins. 1982. ANSWERS (Areal Nonpoint-Source Watershed Environment Response Simulation): User's Manual. U.S. EPA-905/9-82-001, Chicago, Illinois, 54 pp.

Beasley, D.B., L.F. Huggins, and E.J. Monke. 1980. ANSWERS: A model for watershed planning. Trans. Am. Soc. Agric. Eng. 23:938-946.

Beven, K.J. 1977. Experiments with a finite element model of hillslope hydrology. In H.J. Morel-Seytoux, ed., Surface and Subsurface Hydrology, pp. 37-51. Water Resources Publications, Fort Collins, Colorado.

Beven, K.J. 1981. Kinematic subsurface stormflow. Water Resour. Res. 17:1419-1424.

Beven, K.J. 1982. On subsurface stormflow: Predictions with simple kinematic theory for saturated and unsaturated flows. Water Resour. Res. 18:1627-1633.

Beven, K.J., and M.J. Kirkby. 1977. Considerations in the development and validation of a simple physically-based, variable contributing area model of catchment hydrology. In H.J. Morel-Seytoux, ed., Surface and Subsurface Hydrology, pp. 23-36. Water Resources Publications, Fort Collins, Colorado.

Beven, K.J., and M.J. Kirkby. 1979. A physically based variable contributing area model of basin hydrology. Hydrol. Sci. Bull. 24:43-69.

Beven, K.J, and E.F. Wood. 1983. Catchment geomorphology and the dynamics of contributing areas. J. Hydrol. 65:139-158.

Bobba, A.G., D.C.L. Lam, D.S. Jeffries, and others. 1985. Modelling the hydrological regimes in acidified watersheds. In Proceedings for the Muskoka Conference, preprint.

Burch, G.J., C.J. Barnes, I.D. Moore, and others. 1988. Detection and prediction of sediment sources in catchments: Use of Be-7 and Cs-137. Proc. Hydrol. and Water Resour. Symposium 1988, Inst. Engr. Australia, pp. 146-151.

Burch, G.J., R.K. Bath, I.D. Moore, and E.M. O'Loughlin. 1987. Comparative hydrological behavior of forested and cleared catchments in southeastern Australia. J. Hydrol. 90:19-42.

Chapman, T.G. 1980. Modeling groundwater flow over sloping beds. Water Resour. Res. 16:1114-1118.

Childs, E.C. 1971. Drainage of groundwater resting on a sloping bed. Water Resour. Res. 7:1256-1263.

Christophersen, N., S. Kjaernsrod, and A. Rodhe. 1985. Preliminary evaluation of flow patterns in the Birkenes Catchment using natural oxygen-18 as a tracer. In I. Johansson, ed., Hydrological and Hydrogeochemical Mechanisms and Model Approaches to the Acidification of Ecological Systems. Nordic Hydrological Program Report 10, 29-40.

Craig, H. 1961. Standard for reporting concentrations of deuterium and oxygen-18 in natural waters. Science. 133:1833-1834.

Crawford, N.H., and R.K. Linsley. 1966. Digital Simulation in Hydrology: Stanford Watershed Model IV. Technical Report No. 39, Dept. of Civil Engineering, Stanford University, 210 pp.

Davis, L.A., and S.P. Neuman. 1983. Documentation and user's guide: UNSAT2, variably saturated flow model (including 4 example problems). Final Report. NUREG/CR-3390 WWL/TM-1791-1. Water, Waste & Lane, Inc., Ft. Collins, Colorado.

DeCoursey, D.G. 1985. Mathematical models for nonpoint water pollution control. J. Soil Water Conserv. 40:408-413.

Dickinson, W.T., and G.J. Wall. 1977. The relationship between source-area erosion and sediment yield. Hydrol. Sci. Bull. 20:527-530.

Dillaha, T.A., and D.B. Beasley. 1983. Distributed parameter modelling of sediment movement and particle size distributions. Trans. Am. Soc. Agric. Eng. 26:1766-1772, 1777.

Donigian, A.S., D.C. Beyerlein, H.H. Davis, and N.H. Crawford. 1977. Agricultural Runoff Management (ARM) model version II: Refinements and testing. EPA-600/3-77-098, U.S. EPA, Washington, D.C.

Donigian, A.S., and N.H. Crawford. 1976. Modelling nonpoint-source pollution from the land surface. EPA-600/3-76-083 U.S. EPA, Washington, D.C., 279 pp.

Donigian, A.S., and H.H. Davis. 1978. User's manual for Agricultural Runoff Management (ARM) model. EPA-600/3-78-080, U.S. EPA, Washington, D.C., 163 pp.

Dunne, T. 1978. Field studies of hillslope flow processes. In M.J. Kirkby, ed., Hillslope Hydrology, pp. 227-293, Wiley-Interscience, New York.

Dunne, T. 1983. Relation of field studies and modeling in the prediction of storm runoff. In I. Rodrigues-Iturbe and V.K. Gupta, eds., Scale Problems in Hydrology. J. Hydrol. 65:25-48.
Faure, G. 1986. Principles of Isotope Geology (2nd ed.). 589 p. John Wiley & Sons, New York.

Federer, C.A., and D. Lash. 1978. BROOK: A hydrologic simulation model for eastern forests. Water Resour. Res. Center, University of New Hampshire, Durham, N.H., Res. Rept. No. 19, 84 pp.

Foster, G.R., L.J. Lane, D.L. Schertz, and others. 1987. User requirements: USDA-water erosion prediction project (WEPP). Paper No. 87-2539, Am. Soc. Agr. Engr., St. Joseph, Michigan, 6 pp.

Freeze, R.A. 1971. Three-dimensional, transient, saturated-unsaturated flow in a groundwater basin. Water Resour. Res. 7:347-366.

Freeze, R.A. 1972a. Role of subsurface flow in generating surface runoff.
1. Base flow contributions to channel flow. Water Resour. Res. 8:609-623.

Freeze, R.A. 1972b. Role of subsurface flow in generating surface runoff.
2. Upstream source areas. Water Resour. Res. 8:1271-1283.

Freeze, R.A. 1974. Streamflow generation. Rev. Geophys. Space Phys., 12: 627-647.

Freeze, R.A. 1978. Mathematical models of hillslope hydrology. In M.J. Kirkby, ed., Hillslope Hydrology, pp. 177-225, Wiley-Interscience, New York.

Fritz, P., J.A. Cherry, K.U. Weyer, and M.G. Sklash. 1976. Runoff analyses using environmental isotopes and major ions. In Interpretation of Environmental Isotope and Hydrochemical Data in Groundwater Hydrology, I.A.E.A., Vienna, pp. 111-130.

Fritz, P. and J.Ch. Fontes. 1980. Introduction. In P. Fritz, and J.Ch. Fontes, eds., Handbook of Environmental Isotope Geochemistry, volume 1. The Terrestrial Environment, A., pp. 1-19, Elsevier, Amsterdam.

Gat, J.R. 1980. The isotopes of hydrogen and oxygen in precipitation. In P. Fritz, and J.Ch. Fontes, eds., Handbook of Environmental Isotope Geochemistry, volume 1. The Terrestrial Environment, A., pp. 21-47, Elsevier, Amsterdam.

Gat, J.R., and R. Gonfiantini, eds. 1981. Stable Isotope Hydrology. Deuterium and Oxygen-18 in the Water Cycle. I.A.E.A., Vienna, 339 pp.

Gilley, J.E., S.C. Finkner, J.R. Simanton, and G.A. Weesies. 1987. Hydraulic roughness coefficients for upland areas. Paper No. 87-2573, Am. Soc. Agric. Engr., St. Joseph, Michigan, 12 pp.

Henderson, F.M., and R.A. Wooding. 1964. Overland flow and groundwater flow from a steady rainfall of finite duration. J. Geophys. Res. 69:1531-1540.

Hewlett, J.D., and A.R. Hibbert. 1967. Factors affecting the response of small watersheds to precipitation in humid areas. In W.E. Sopper, and H.W. Lull, eds., International Symposium on Forest Hydrology, pp.275-290, Permagon, New York.

Hoefs, J. 1987. Stable Isotope Geochemistry (3rd ed.), Springer-Verlag, New York, 240 pp.

Hurley, D.G., and G. Pantelis. 1985. Unsaturated and saturated flow through a thin porous layer on a hillslope. Water Resour. Res. 21:821-824.

Kennedy, V.C., C. Kendall, G.W. Zellweger, and others. 1986. Determination of the components of stormflow using water chemistry and environmental isotopes, Mattole River Basin, California. J. Hydrol. 84:107-140.

Knisel, W.G. ed. 1980. CREAMS: A field scale model for Chemicals, Runoff, and Erosion from Agricultural Management Systems. USDA Conserv. Res. Rep. No. 26, Sci. and Ed. Admin., 643 pp.

Lane, L.J., G.R. Foster, and A.D. Nicks. 1987. Use of fundamental erosion mechanics in erosion prediction. Paper No. 87-2540, Am. Soc. Agric. Engr., St. Joseph, Michigan, 10 pp.

Leytham, K.M., and R.C. Johanson. 1979. Watershed erosion and sediment transport model. EPA-600/3-79-028, U.S. EPA, Washington, D.C., 357 pp.

Lomenick, T.F., and T. Tamura. 1965. Naturally occurring fixation of cesium-137 on sediments of lacustrine origin. Proc. Soil Sci. Soc. Am. 29:383-387.

Maloszewski, P., W. Rauert, W. Stichler, and A. Hermann. 1983. Application of flow models in an alpine catchment area using tritium and deuterium data. J. Hydrol., 66:319-330.

Mohsenisarvi, M. 1981. Forecasting subsurface water flow and storage on forested slopes using a finite element model. Ph.D. Thesis, Univ. of Idaho, Univ. Microfilms International, Ann Arbor, Michigan, 162 pp.

Moore, I.D., and G.J. Burch. 1986. Modelling erosion and deposition: topographic effects. Trans. Am. Soc. Agric. Engr. 29:1624-1630, 1640.

Moore, I.D., G.J. Burch, and D.H. Mackenzie. 1988a. Topographic effects on the distribution of surface soil water and the location of ephemeral gullies. Trans. Am. Soc. Agric. Engr. 31:1098-1107.

Moore, I.D., G.J. Burch, and P.J. Wallbrink. 1986a. Preferential flow and hydraulic conductivity of forest soils. Soil Sci. Soc. Am. J. 50:876-881.

Moore, I.D., G.B. Coltharp, and P.G. Sloan. 1983. Predicting runoff from small Appalachian watersheds. Trans. Kentucky Academy Sci. 44:135-145.

Moore, I.D., S.M. Mackay, P.J. Wallbrink, and others. 1986b. Hydrologic characteristics and modelling of a small forested catchment in southern New South Wales: Pre-logging condition. J. Hydrol. 83:307-335.

Moore, I.D., E.M. O'Loughlin, and G.J. Burch. 1988b. A contour-based topographic model for hydrological and ecological applications. Earth Surface Processes and Landforms 13:305-320.

Mosley, M.P. 1979. Streamflow generation in a forested watershed, New Zealand. Water Resour. Res. 15:795-806.

Mosley, M.P. 1982. Subsurface flow velocities through selected forest soils, South Island, New Zealand. J. Hydrol. 55:65-92.

Neuman, S.P. 1973. Saturated-unsaturated seepage by finite elements. J. Hydraul. Div., Proc. Am. Soc. Civil Engr. 99:2233-2250.

Nieber, J.L. 1982. Hillslope soil moisture flow, approximation by a one-dimensional formulation. Paper No. 82-2026, Am. Soc. Agric. Engr. St. Joseph, Michigan.

Nieber, J.L., and M.F. Walter. 1981. Two-dimensional soil moisture flow in a sloping rectangular region: experimental and numerical studies. Water Resour. Res. 17:1722-1730.

O'Loughlin, E.M. 1981. Saturation regions in catchments a and their relations to soil and topographic properties. J. Hydrol. 53:229-246.

O'Loughlin, E.M. 1986. Prediction of surface saturation zones in natural catchments by topographic analysis. Water Resour. Res. 22:794-804.

Payne. B.R., and E. Halevy. 1968. Environmental isotopes. In Guidebook on Nuclear Techniques in Hydrology, I.A.E.A. Technical Report Series No. 91, I.A.E.A., Vienna pp. 5-14.

Pearce, A.J., and A.I. McKerchar. 1979. Upstream generation of storm runoff. In D.L. Murray, and P. Ackroyd eds., Physical Hydrology - New Zealand Experience, pp. 165-192, New Zealand Hydrological Society, Wellington, pp. 165-192.

Pearce, A.J., M.K. Stewart, and M.G. Sklash. 1986. Storm runoff generation in humid headwater catchments, I: Where does the water come from? Water Resour. Res. 22:1263-1272.

Reeves, M., and J.O Duguid. 1975. Water movement through saturated-unsaturated porous media: a finite element Galerkin model. Oak Ridge National Laboratory, Oak Ridge, Tennessee. NTIS: ORNL-4927, 232 pp.

Ritchie, J.C., J.A. Spraberry, and J.R. McHenry. 1974. Estimating soil erosion from the redistribution of fallout Cs-137. Soc. Sci. Soc. Am. J., 38:137-139.

Rodhe, A. 1987. The origin of streamwater traced by oxygen-18. Uppsala Univ., Dept. Phys. Geogr., Div. Hydrol., Report Series A, No. 41, 260 pp. Appendix 73 pp.

Rose, C.W., J.Y. Parlange, G.C. Sander, and others. 1983a. Kinematic flow approximation to runoff on a plane: An approximate analytic solution. J. Hydrol. 62:263-269.

Rose, C.W., J.R. Williams, G.C. Sander, and D.A. Barry. 1983b. A mathematical model of soil erosion and deposition processes: I. Theory for a plane land element. Soil Sci. Soc. Am. J. 47:991-995.

Rose, C.W., J.R. Williams, G.C. Sander, and D.A. Barry. 1983c. A mathematical model of soil erosion and deposition processes: II. Application to data from an arid-zone catchment. Soil Sci. Soc. Am. J. 47:996-1000.

Sisson, J.B., A.H. Ferguson, and M.T. van Genuchten. 1980. Simple method for predicting drainage from field plots. Soil Sci. Soc. Am. J. 44:1147-1152.

Sklash, M.G. 1978. The role of groundwater in storm and snowmelt runoff. Ph.D. thesis, University of Waterloo, 261 pp.

Sklash, M.G. and Farvolden, R.N. 1979. The role of groundwater in storm runoff. J. Hydrol. 43:45-65.

Sklash, M.G., M.K. Stewart, and A.J. Pearce. 1986. Storm runoff generation in humid headwater catchments, II: A case study of hillslope and low order stream response. Water Resour. Res. 22:1273-1282.

Sloan, P.G., and I.D. Moore. 1983. Modelling subsurface stormflow on steeply sloping watersheds. Water Resour. Res. 20:1815-1822.

Smith, R.E., and R.H.B. Hebbert. 1983. Mathematical simulation of inter-dependent surface and subsurface hydrologic processes. Water Resour. Res. 19:987-1001.

Smith, R.E., and D.A. Woolhiser. 1971. Overland flow on an infiltrating surface. Water Resour. Res. 7:899-913.

Stagnitti, F., M.B. Parlange, T.S. Steenhuis, and J.Y. Parlange. 1986. Drainage from a uniform soil layer on a hillslope. Water Resour. Res. 22:631-634.

Stephenson, G.R., and R.A. Freeze. 1974. Mathematical simulation of subsurface flow contributions to snowmelt runoff, Reynolds' Creek Watershed, Idaho. Water Resour. Res. 10:284-298.

Storm, D.E. 1986. Modelling phosphorus transport in surface runoff from agricultural watersheds for nonpoint-source pollution assessment. M.S. Thesis, Virginia Polytechnic Inst. and State Univ., Blacksburg, Va, 202 pp.

Thomas, D.L., and D.B. Beasley. 1986a. A physically-based forest hydrology model. I: Development and sensitivity of components. Trans. Am. Soc. Agric. Engr. 29:962-972.

Thomas, D.L., and D.B. Beasley. 1986b. A physically-based forest hydrology model. II: Evaluation under natural conditions. Trans. Am. Soc. Agric. Engr. 29:973-981.

Thorne, C.R., L.W. Zevenbergen, F.H. Grissinger, and J.B. Murphey. 1986. Ephemeral gullies as sources of sediment. Proc. Fourth Federal Interagency Sed. Conf., Las Vegas, Nev., March 24-27, vol. 1, pp. 3.152-3.161.

Troendle, C.A. 1979. A variable source area model for storm flow prediction on first order forested watershed. Ph.D. Thesis, Univ. Georgia, 115 pp.

Turner, J.V., D.K. Macpherson, and R.A. Stokes. 1987. The mechanisms of catchment flow processes using natural variations in deuterium and oxygen-18. In A.J. Peck, and D.R. Williamson, eds., Hydrology and Salinity in the Collie River Basin, Western Australia, 94:143-162, J. Hydrol.

Wall, G.J., W.T. Dickinson, and L.J.P. van Vliet. 1979. Agricultural sources of fluvial suspended sediments. Prog. Water. Technol. 11:481-499.

Ward, R.C. 1984. On the response to precipitation of headwater streams in humid areas. J. Hydrol. 74:171-189.

Williams, J.R. 1975a. Sediment yield prediction with Universal Equation using runoff energy factor. In Present and Prospective Technology for Predicting Sediment Yields and Sources. ARS-S-40, USDA-Agric. Res. Service, pp. 244-252.

Williams, J.R. 1975b. Sediment routing for agricultural watersheds. Water Resour. Bull. 11:965-974.

Williams, J.R. 1978. A sedimentgraph model based on an instantaneous unit sedimentgraph. Water Resour. Res. 14:659-664.

Wilson, B.N., B.J. Barfield, I.D. Moore, and R.C. Warner. 1984b. A hydrology and sedimentology watershed model. Part II: Sedimentology component. Trans. Am. Soc. Agric. Engr. 27:1378-1384.

Wilson, B.N., B.J. Barfield, A.D. Ward, and I.D. Moore. 1984a. A hydrology and sedimentology watershed model. Part I: Operational format and hydrologic component. Trans. Am. Soc. Agric. Engr. 27:1378-1384.

Wischmeier, W.H., and D.D. Smith. 1978. Predicting rainfall erosion losses. Agric. Handb. 537, USDA-Sci. and Ed. Admin., 58 pp.

Young, R.A., C.A. Onstad, D.D. Bosch, and W.P. Anderson. 1987. AGNPS, agricultural nonpoint-source pollution model; a large watershed analysis tool. USDA-Agr. Res. Serv., Conserv. Res. Rep. 35, 77 pp.

Young, R.A., C.A. Onstad, D.D. Bosch, and W.P. Anderson. 1989. AGNPS: A nonpoint-source pollution model for evaluating agricultural watersheds. J. Soil and Water Conserv. 2:168-173.

Young, J.A., and W.B. Silker. 1980. Aerosol deposition velocities on the Pacific and Atlantic oceans calculated from Be-7 measurements. Earth and Planetary Science Letters, 50:92-104.

MULTIMEDIA MODELING OF TOXIC CHEMICALS

Y. Onishi[1], L. Shuyler[2] and Yoram Cohen[3]

ABSTRACT

Many persistent toxic chemicals released to the environment undergo complex interactions in multiple environmental media: land surface, groundwater, surface water, and air. Mathematical multimedia models can integrate many transport and fate mechanisms into a single framework to improve the accuracy of environmental exposure and risk assessment.

INTRODUCTION

The release of substances from point and nonpoint sources into the environment and the subsequent effects on biota and human health are major environmental issues. Considerable efforts are being made to reduce the release of toxic contaminants to the environment. Because lowering pollutant concentration levels is costly, optimal environmental management is best achieved via cost-benefit analysis. Thus, decision makers must have a sound basis on which to rely when assessing impact.

Many toxic contaminants are persistent and undergo complex interactions in multiple environmental media such as land surface, groundwater, surface water, and air. Thus, optimal environmental management requires estimation of toxic chemical distribution and fate in multiple environmental media. Mathematical models, especially multimedia models that include groundwater, surface water, leaching, and atmospheric components, can be used to integrate many complex mechanisms controlling the transport and fate of toxic chemicals in the environment into a single framework to improve the accuracy of environmental exposure and risk analysis.

MULTIMEDIA TRANSPORT AND FATE PROCESSES AND MECHANISMS

Contaminants released to the environment are spread and dissipated by the processes of a) transport, b) degradation and decay, c) transformation, d) intermedia transfer among overland, groundwater, surface water, and atmosphere pathways, and e) biological uptake (fig. 1).

The primary pathways for any medium begin at the contaminant source. These pathways generally present problems of local-scale and acute environmental effects, with high concentrations of contaminants emanating from the source. A notable exception is the acid rain problem, which is regional in scale.

Secondary pathways are those in which the contaminant is transported from one environmental pathway to another. These secondary pathways generally represent regional-scale problems of chronic environmental effects from lower concentrations of contaminant because contaminant dispersion and dilution have already occurred in the primary pathway.

[1]Y. Onishi, Technical Leader, Battelle, Pacific Northwest Laboratories, Richland, WA.

[2]Lynn Shuyler, U.S. Environmental Protection Agency, Region III, Annapolis, MD.

[3]Yoram Cohen, Associate Professor, University of California at Los Angeles, Los Angeles, CA.

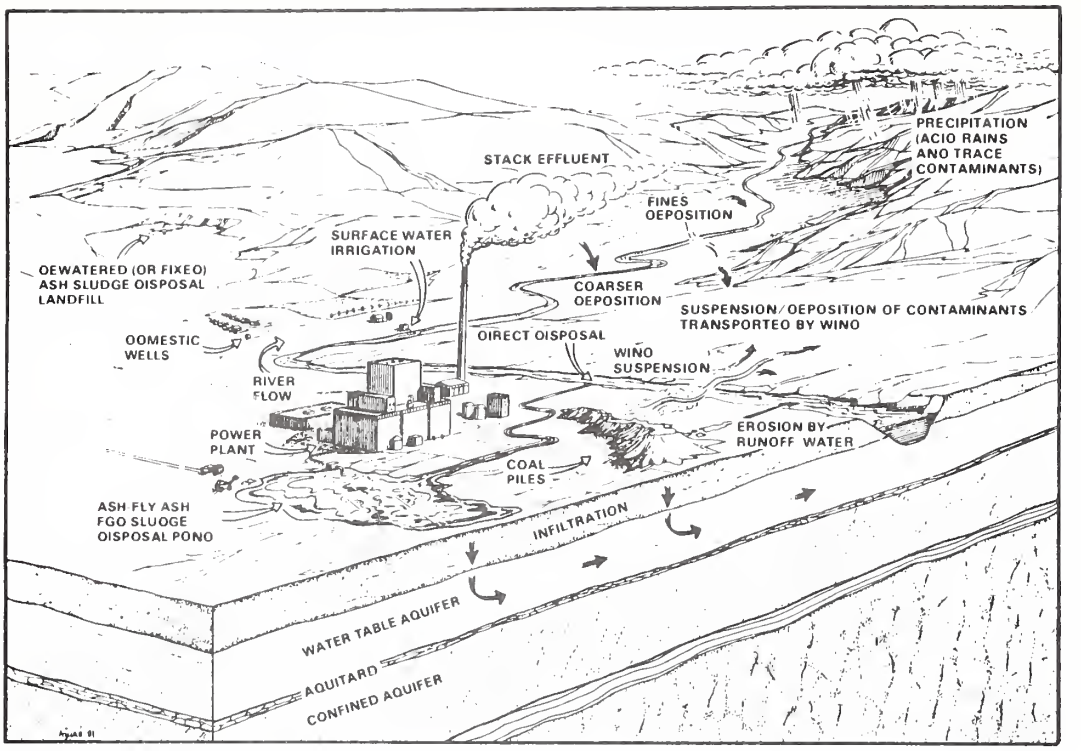


Figure 1.
Various pathways and potential sources for environmental contamination.

The interactions of various pathways and the linkage of these pathways to humans are illustrated in figure 2. Because many contaminants migrate from one pathway to others, the associated effects to humans and environments can derive from more than one environmental medium. After pesticide applications, for example, pesticide losses from agricultural lands can occur through infiltration to groundwater; runoff and soil erosion over land surface and to receiving surface water; resuspension and volatilization to the atmosphere; degradation (microbial, chemical, and photochemical), and uptake by plants and animals. Thus, the multimedia pathways must be assessed to evaluate migration and potential risks of pesticides. The example in figure 3 shows potential sequential primary and secondary pathways.

Another example of the multimedia nature of environmental pollution is acid rain. In this case, SO_x and NO_x when released to the atmosphere are transported downwind and subsequently washed out of the air in rainfall as sulfuric and nitric acids, resulting in acid contamination of agricultural lands, forests, and lakes. The potential sequential pathways to be considered in this case are delineated in figure 4.

MULTIMEDIA MODELING APPROACHES

Various multimedia assessment methodologies have been developed to integrate these transport and fate mechanisms in a logical and consistent framework. Three approaches for representing these complex interactions are

(1) fully coupled, integrated multimedia modeling,
(2) partially coupled, integrated multimedia modeling, and
(3) composite multimedia modeling.

Each of these methodologies has been developed for a particular objective and therefore cannot be applied arbitrarily to meet all assessment conditions and objectives.

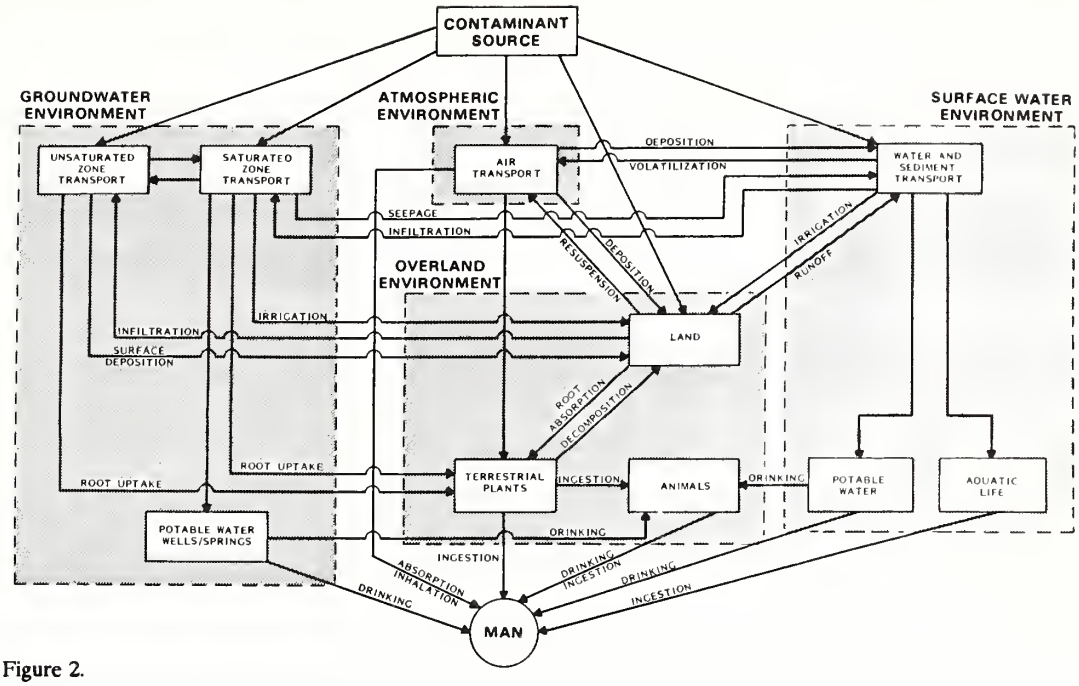


Figure 2.
Interactions among pathways and their effects on humans and the environment.

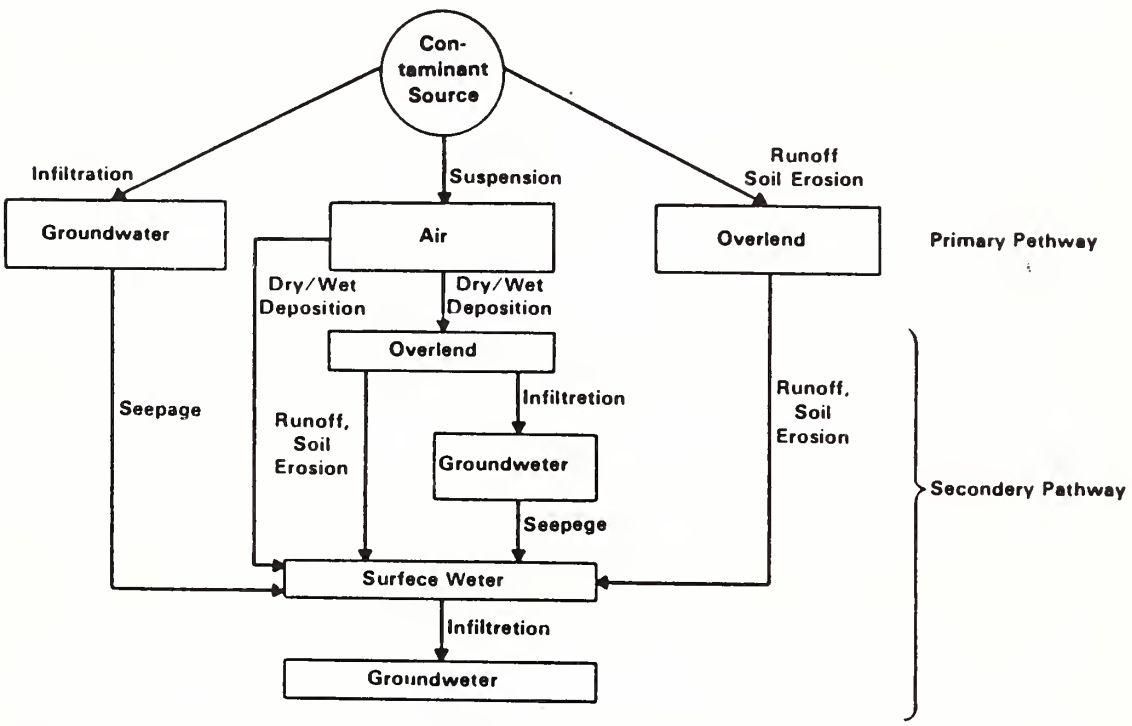


Figure 3.
Sequential primary and secondary pathways for overland releases.

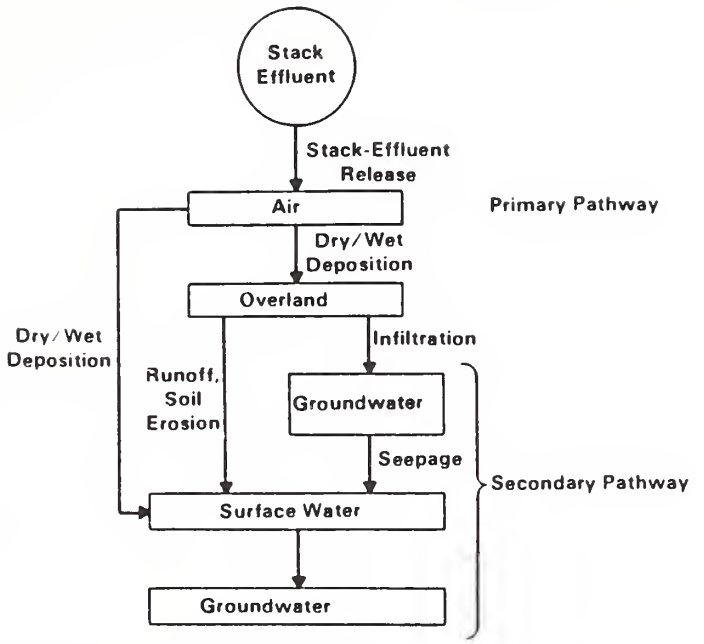


Figure 4.
Sequential primary and secondary pathways for atmospheric release.

Fully Coupled, Integrated Multimedia Modeling

Fully Coupled, Integrated Multimedia Modeling Approach

Fully coupled, integrated multimedia modeling integrates the submodels for each environmental compartment (e.g., soil, groundwater, surface water, and air) as parts of an overall multimedia model. Coupling and integration thus allow feedback interactions among the environmental compartments by solving the governing equations of the various pathway models simultaneously. Because of the complexity of transport processes and the spatial variability in media properties, such a model can become extremely large and cumbersome to implement in a unified manner. To make such an approach practical, then, some simplifications are introduced to reduce the required input data and the required code size and run times. In the first level of simplification, all of the environmental compartments are assumed to be uniform. Thus, spatial transport within each environmental medium is neglected, and contaminant transport among the well-mixed compartments is described by a variety of cross-media transport processes. In the second level of complexity, some compartments are allowed to be nonuniform while others remain uniform. This latter hybrid, fully coupled, integrated modeling approach is designed to retain the practical simplicity of the models and yet to introduce the inherent spatial variability in some compartments (e.g., soil and bottom sediment of surface water).

Multimedia compartmental models: uniform compartments. Elimination of spatial resolution of contaminant distribution within each pathway allows reduction of the governing transport equations from a system of partial differential equations in time and space to a system of ordinary differential equations in time. For a given environmental pathway, the governing equation is (Cohen and Ryan 1985; Mayer 1988):

$$V_i \frac{dC_i}{dt} = \sum_{\substack{j=1 \\ j\neq 1}}^{N} A_{ij} K_{ij} (C^*_{ij} - C_i) + R_i + S_i + \sum_{j=1}^{N} (Q_{ij}C_j - Q_{ij}C_i) \quad i=1,...N \qquad [1]$$

where A_{ij} = interface area between i-th and j-th environmental media [L^2]
C_i = contaminant concentration in i-th environmental medium [ML^{-3}]
C_{ij} = interface concentration in i-th medium, which is in equilibrium with the bulk concentration of the j-th medium [ML^{-3}]
K_{ij} = i-th side overall mass transfer coefficient between i-th and j-th media [LT^{-1}]
N = number of environmental media
Q_{ij} = volumetric flow rate from i-th to j-th medium [L^3T^{-1}]
R_i = contaminant production rate in i-th environmental medium from chemical degradation and decay, volatilization, transformation, biological degradation, etc. [MT^{-1}]
V_i = volume of i-th environmental medium [L^3]
S_i = contaminant source strength, such as contaminant release rate to i-th medium [MT^{-1}]
t = time [T].

In the steady state, the time-dependent term will be dropped from equation 1, thus further simplifying the governing equations from ordinary differential to algebraic equations.

The significant simplification (especially the elimination of transport processes - advection and dispersion - within each compartment) imposed on these models means these fully coupled, integrated multimedia models with uniform environmental compartments require modest input data and are very easy to run. Thus, despite the loss in spatial resolution of a contaminant, fully coupled, integrated multimedia models may be useful as a screening methodology to evaluate overall environmental responses to contaminant releases. These models are limited to screening analysis, however, because they cannot provide accurate estimates of contaminant concentrations in applications in which spatial variability is important.

Fully coupled, integrated multimedia modeling approach: combination of nonuniform and uniform compartments. Because the transport in some environmental compartments is dominated by slow diffusion processes, highly nonuniform concentration profiles are often established. For example, the soil and bottom sediment compartments are inherently nonuniform (in contaminant concentration) and thus must be treated accordingly. In contrast, large portions of the atmosphere and surface water compartments can be treated to be well mixed. When including compartments where mixing is intense (compared to the soil and sediment compartments, for example), subcompartments can be defined to incorporate some degree of stratification.

Cohen and coworkers (Mayer 1988) recently demonstrated this modeling approach with the Spatial Multimedia Compartmental (SMCM) model. The SMCM model incorporates nonuniform compartments such as soil and sediments by describing contaminant transport in these compartments via the one-dimensional convection-diffusion model:

$$\frac{\partial C}{\partial t} + u\frac{\partial C_i}{\partial x} = \frac{\partial}{\partial x}\left(D_i \frac{\partial C_i}{\partial x}\right) - R_i \qquad [2]$$

where C_i = overall concentration in the soil matrix [ML^{-3}]
D_i = soil matrix effective diffusion coefficient (including dispersion) [L^2T^{-1}]
u = convective velocity [LT^{-1}]
R_i = any chemical or biochemical reactions that may consume or produce the contaminant under consideration [$ML^{-3}T^{-1}$].

The compartmental description of the SMCM model (figure 5) thus requires the simultaneous solution of a minimum of four ordinary differential equations [equation 1 for air, surface water, suspended solids, and biota] and two partial differential equations [equation 2 for soil and bottom

sediment]. Some spatial variability can be introduced into the air and surface water compartments by further dividing these compartments into subcompartments with different mixing characteristics (temperature, size, etc.). Finally, as with other spatial models, the SMCM model accounts for runoff, infiltration, and soil drying, as well as rain scavenging of pollutants.

Fully Coupled, Integrated Multimedia Models

Uniform compartments. Typical models in this category are the Multimedia Compartment (MCM) model (Cohen 1981), the GEOTOX model (McKone 1985), and fugacity models (Mackay and Paterson 1982). The MCM is a six-compartment model with each compartment representing uniformly mixed environmental pathways of air, soil, surface water, suspended solids, aquatic biota, and surface water bottom sediment. The model simultaneously solves six sets of governing equations for each compartment (pathway), that is, equation 1 with a known initial contaminant concentration in each compartment as initial conditions. The model was applied to estimate environmental distributions of benz(a)pyrene [b(a)p] in southeast Ohio and trichloroethylene (TCE) in San Diego, California.

The GEOTOX model incorporates exposure assessment and risk assessment into a multimedia framework to evaluate health risks of trace elements, radionuclides, and organic chemicals (McKone and Kastenburg 1985). The model includes seven compartments, representing air, land biota, upper soil layer, lower soil layer, groundwater zone, surface water, and surface water bottom sediments. With appropriate multimedia transfer rates, the seven ordinary differential equations expressing contaminant mass balance are numerically solved simultaneously to obtain contaminant distributions in these seven pathways. These computed concentrations are then used to determine potential health risks associated with contaminant exposure.

Mackay and associates developed a series of fugacity multimedia models to estimate contaminant distributions in air, soil, surface water, suspended solids, biota, and surface water bottom sediment (Mackay 1987). Fugacity is a thermodynamic measure of the escaping tendency of the substance within a system (Mayer 1988). In the fugacity models, contaminant concentration (C) is replaced by fugacity in the following relationship.

$$C = z\,f \qquad [3]$$

where f = fugacity
z = fugacity capacity.

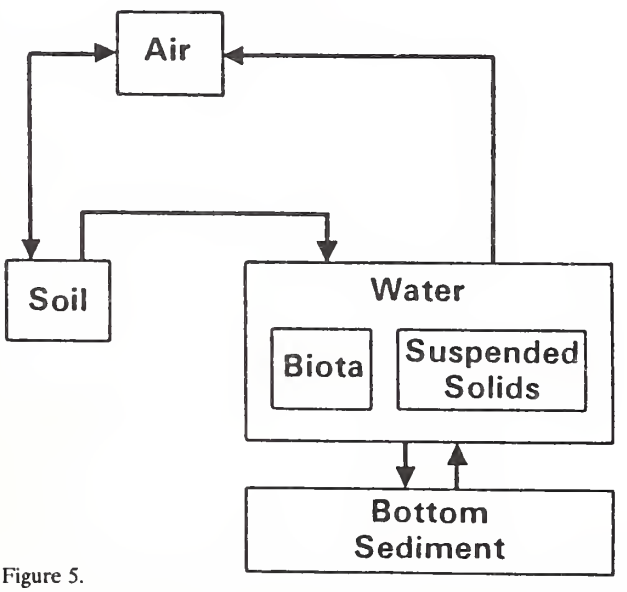


Figure 5.
Configuration of the SMCM multimedia compartmental model (Mayer 1988).

The advantage of the use of fugacity is to allow the transport parameters (e.g., mass transfer coefficients and partition coefficients) to have the same units across the various multimedia (Mayer 1988).

Combined nonuniform and uniform compartments. An example of models in this category is the Spatial Multimedia Compartmental Model (SMCM) developed by Cohen and coworkers (Mayer 1988). The SMCM model was applied to estimate the distribution of benzene and trichloroethylene (TCE) in the Santa Clara Valley, California. To make the model convenient for screening applications, the SMCM model incorporates user-friendly full-screen user data input screens with on-line help windows. Various default parameter values are provided but these can be easily replaced by the user. Thus, the SMCM model is easy to operate and can be used to investigate a variety of scenarios.

Partially Coupled, Integrated Multimedia Modeling

Partially Coupled, Integrated Multimedia Modeling Approach

In the partially coupled, integrated multimedia modeling approach, each environmental medium is represented by a submodel internally connected to form a single multimedia model. However, models in the partially integrated approach are unlike those using the fully integrated approach in that the governing equations are solved sequentially rather than simultaneously. Interfacing and information transfer between environmental pathways are usually managed by a central executive program. The specific environmental pathway submodels are selected a priori by the original model developer. These models therefore are somewhat limited in applicability to broader site and release scenario conditions, but allow consistent and unified descriptions of the environmental systems to achieve consistent applications from site to site (Whelan et al. 1987).

Partially coupled, integrated multimedia models solve pathway submodels sequentially. Each submodel handles transport and fate phenomena in a more sophisticated manner than the fully coupled, integrated models, but is still simple enough to be incorporated in a single model. These submodels usually solve spatial distributions of contaminants within each environmental pathway. As compared to fully coupled, integrated multimedia models, these partially coupled, integrated multimedia models require more input data and computational resources and may provide more accurate results. However, they are more easily applied and potentially less accurate in their predictions than composite multimedia models, which will be discussed later.

Partially Coupled, Integrated Multimedia Models

Representative models in this category include the Water Transport Model (Fletcher and Dodson 1971), the Hydrologic Simulation Program in FORTRAN (HSPF) mode (Johanson et al. 1980; Donigian et al. 1983), the Air Land Water Analysis System (ALWAS) (Tucker et al. 1984), the Unified Transport model (UTM) (Patterson 1986), and the Remedial Action Priority System (RAPS) (Whelan et al. 1987).

The model developed by Fletcher and Dodson (1971) is an example of a series of multimedia models used to estimate radiation doses to humans from nuclear facilities (Onishi et al. 1981b). The overall model structure of the model is shown in figure 6, and a detailed water transport portion of the model is shown in figure 7. As shown in these figures, the liquid pathway model includes dissolved and particulate radionuclide transport and sediment transport. Options include input from numerous point sources and air deposition, and from overland flow, flood deposition, and groundwater transport. The liquid pathway model is an unsteady, one-dimensional code that calculates temporal and spatial (longitudinal) distributions of dissolved radionuclide concentration as well as concentration of radionuclides attached to suspended and bottom sediments. The transport model can also account for the effects of flow stratification and sediment trapping in

reservoirs. The dissolved radionuclide concentration at a given location is found by applying the mass conservation equation with radioactive decay:

$$C_{x,t} = \frac{1}{Q_{x,t}} Q_{(x-\Delta x,\, t-\Delta t)} \; C_{(x-\Delta x,\, t-\Delta t)} \; e^{-\lambda \Delta t} + \sum_{i=1}^{n} Q_i C_i \qquad [4]$$

where $C_{x,t}$ = dissolved radionuclide concentration at location x and time t [ML^{-3}]
C_i = dissolved radionuclide concentration of tributary [ML^{-3}]
$Q_{x,t}$ = flow rate at location x and time t [L^3T^{-1}]
Q_i = tributary flow rate [L^3T^{-1}]
λ = radionuclide decay [T^{-1}].

The sediment transport rate, S [MT^{-1}], is found analytically from the following equation:

$$S = A\,Q^b \qquad [5]$$

where Q is the flow rate [L^3T^{-1}] and a and b are constants that must be estimated for each sediment size range. The concentration of radionuclides attached to sediment is calculated from the known dissolved radionuclide concentration and the distribution coefficient value, K_d [L^3M^{-1}]. Because the amount of radionuclide adsorbed to the sediment is not subtracted from the dissolved concentration, strictly speaking the mass conservation in a stream reach is not satisfied.

The Remedial Action Priority System (RAPS) methodology uses empirically, analytically, and semianalytically based mathematical algorithms and a pathway analysis to estimate the potential exposure levels in the environment and associated risks based on limited site information (Whelan et al. 1987). The RAPS methodology considers the following four transport pathways (fig. 8): subsurface (groundwater), overland, surface water, and atmosphere. The transport pathway submodels are systematically integrated with an exposure assessment component that considers the type, time, and duration of exposure and the location and size of the population exposed. The RAPS methodology uses the Hazard Potential Index to estimate relative risks for a comparative purpose among various sites and scenarios rather than to estimate absolute risks.

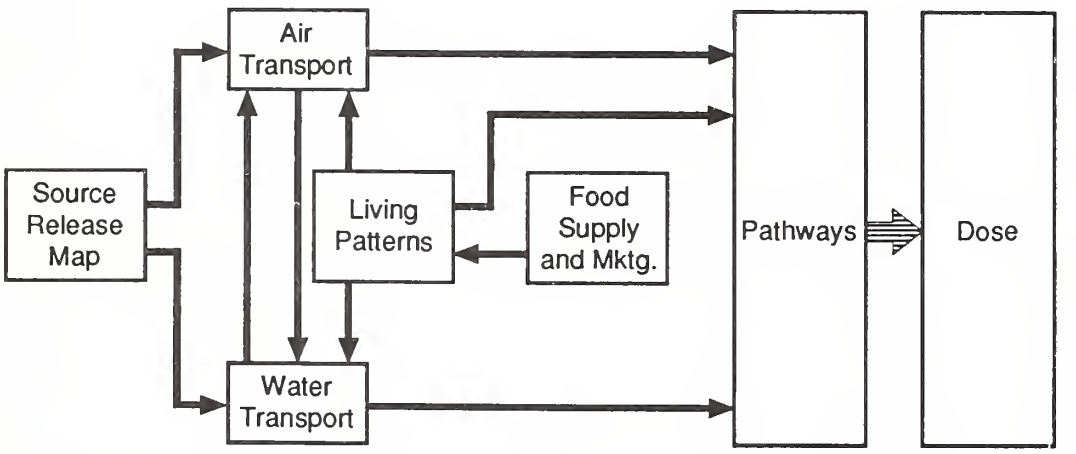


Figure 6.
Overall structure of the model of Fletcher and Dodson (1971).

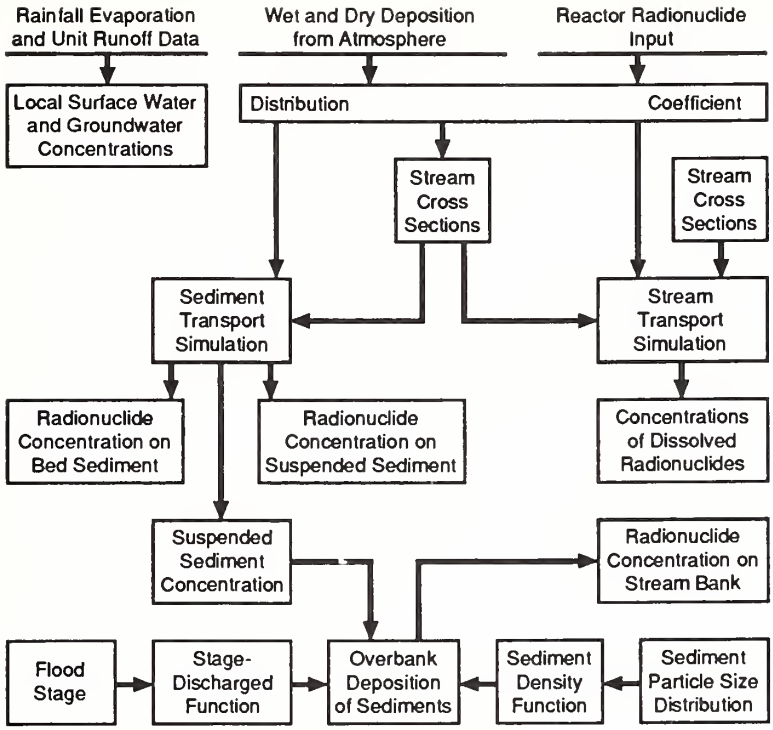


Figure 7.
Water transport model of Fletcher and Dodson (1971).This model expresses the very complicated pathway transport with extremely simplified formulations. Consequently, it is very easy to apply this model, but the results may either overestimate or underestimate the reality. These merits and shortcomings are shared by other man-dose models (Soldat et al. 1974).

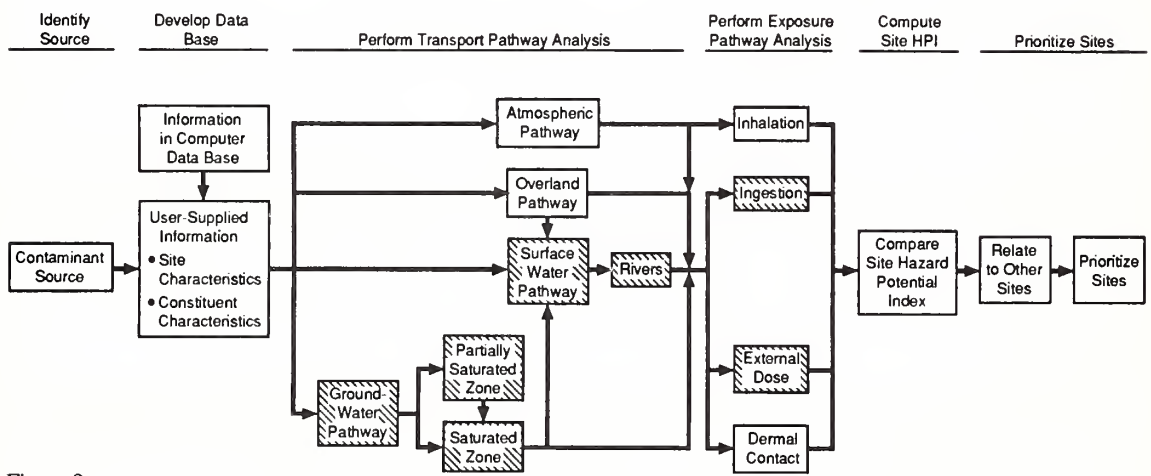


Figure 8.
Interactions between the transport pathways and exposure assessment components of the RAPS Methodology (shaded boxes indicate a potential contaminant transport an exposure route using the RAPS methodology) (Whelan et al. 1987).

Composite Multimedia Modeling

Composite Multimedia Modeling Approach

A composite multimedia methodology assembles appropriate models for various pathways and loosely combines them for multimedia analysis. Because each pathway is represented by an

individual model, more realistic individual pathway models are retained in this case. The user externally couples the models to address the appropriate level of detail dictated by the environmental system and type of assessment required. Therefore, the user conceptualizes the composite modeling scenario whereas conceptualization of the fully or partially coupled scenario is inherent in the design. Composite approach modeling is sequential; codes for individual pathways do not interact directly among themselves. Interfacing and information transfer occurs by assigning the output file from one pathway model to the input file of the next pathway model. Specific feedback among the various environmental pathways, the main justification for the fully coupled multimedia model, is addressed by a model user by analyzing a problem in terms of primary and secondary pathways for each model and the appropriate interactions of these models and their results.

The concepts of the sequential composite multimedia modeling approach are illustrated in figures 3 and 4 for four main types of chemical releases: stack, pile and landfill, waste disposal pond, and direct liquid discharge to surface water. In this approach, typically all primary pathways for each medium are analyzed in a local model area (except a regional air model dealing with acid rain and others) near the source. The outputs from these primary analyses are then fed forward to more regional models, whose results when composited (superpositioned) with the generally more local results of the primary pathway and analyses produce the concentration or rate-versus-time data required to perform the environmental effects analysis. On rare occasions, a secondary pathway modeling may be repeated to account for any significant feedback from another secondary results pathway. The composite multimedia approach allows each component or code to be replaced as the scenario being modeled changes or as technological advances are made. This approach allows the user to customize frameworks to address specific modeling needs and to allocate resources to optimize the resource requirements of the analysis in relation to the goals of the assessment.

The composite multimedia approach is more flexible overall than the fully or partially coupled integrated approach because only the necessary pathway models are used for a given problem. The composite multimedia methodologies require extensive input data and computer resources, but provide the potentially most accurate predictions among time modeling approaches discussed here. Thus, they are most useful to problematic site assessment requiring accurate predictions of contaminant transport and fate (for example, the site remediation studies), after fully or partially coupled, integrated models and/or field data have revealed potential environmental problems. However, this approach demands of the model user more insight to a specific problem because the user is developing a conceptual model tailored to the problem.

Composite Multimedia Models

Examples of sequential pathway modeling using the composite multimedia approach include the Land Disposal Restriction (LDR) methodology (EPA 1986), the Chemical Migration and Risk Assessment (CMRA) methodology (Onishi et al. 1979, 1981a, 1985; Whelan and Parkhurst 1983), and the Multimedia Contaminant Environmental Exposure Assessment (MCEA) methodology (Onishi et al. 1982a,b; Whelan and Onishi 1983). Bolten et al. (1983) expanded the MCEA methodology to include cost and risk analysis components.

The CMRA methodology predicts the frequency and persistence of toxic contaminants (e.g., pesticides, heavy metals, and other hazardous materials) on land surface and in surface waters by combining continuous simulation of transport of dissolved and sediment-sorbed contaminants and associated risk assessment into a single system. It also predicts acute (lethal) and chronic (sublethal) damage to aquatic biota. The methodology consists of the following components: 1) overland contaminant transport modeling; 2) contaminant transport modeling in surface waters (i.e., rivers, estuaries, coastal waters, and lakes); 3) statistical analysis of contaminant concentrations in surface waters; and 4) risk assessment. Figure 9 illustrates how these components are connected and what types of data are needed. A brief description of each component follows.

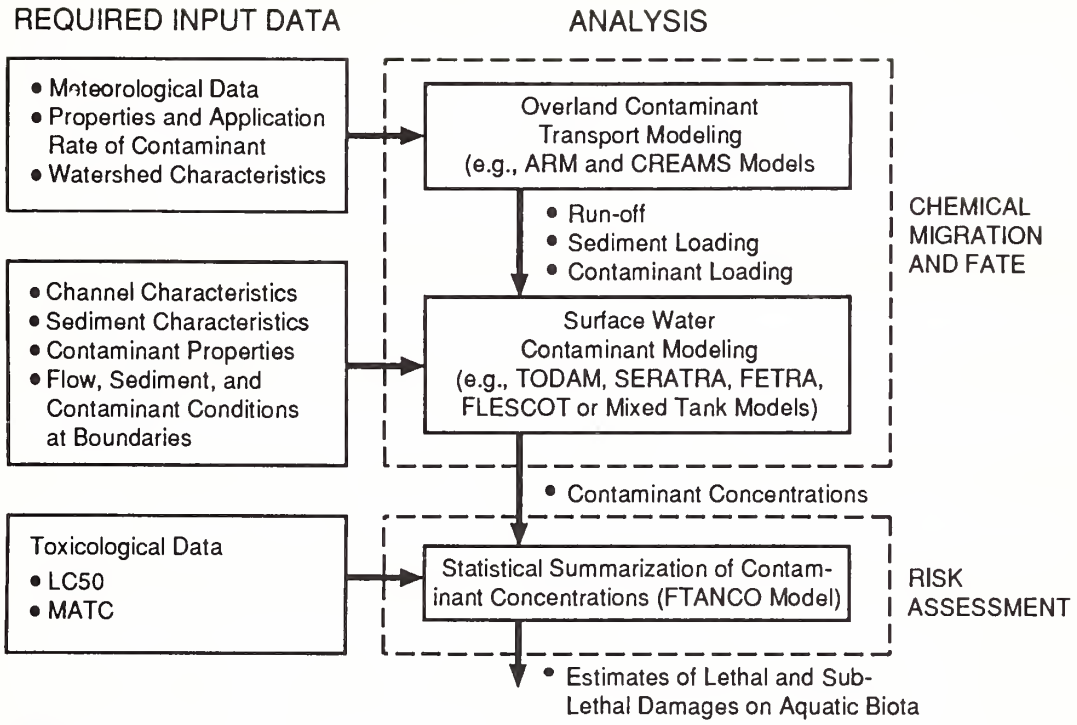


Figure 9.
Chemical migration and risk assessment (CMRA) methodology (Onishi et al. 1985).

The CMRA methodology uses detailed and continuous overland simulation models such as the Agricultural Runoff Management (ARM) model (Donigian and Crawford 1976) and CREAMS (Knisel 1980) for its overland modeling. For example, the ARM model predicts runoff and loadings of sediment and contaminant (both dissolved and particulate) at the edge of the receiving surface water continuously during a simulation period. The CMRA methodology uses one of several unsteady, sediment/contaminant (both dissolved and sediment-sorbed) transport surface water models that include mechanisms for sediment/contaminant interactions, such as adsorption/desorption and transport, deposition, and resuspension of sediment-sorbed contaminants. These models are the one-dimensional models, the Time-Dependent, One-Dimensional Degradation and Migration Model (TODAM) (Onishi et al. 1982a,b) and the Mixed Tank Model (Onishi et al. 1985); the two-dimensional (longitudinal and vertical) Sediment and Radionuclide Transport model (SERATRA) (Onishi and Wise 1979); the two-dimensional (longitudinal and lateral) Finite Element Transport model (FETRA) (Onishi 1981); and the three-dimensional Flow, Energy, Salinity, Sediment and Contaminant Transport model (FLESCOT) (Onishi and Trent 1985).

All these models, except the Mixed Tank Model, have the following submodels to predict sediment and contaminant distribution in a water column and bed:

(1) cohesive and noncohesive sediment transport,
(2) dissolved contaminant transport and degradation, and
(3) particulate contaminant transport associated with cohesive and noncohesive sediments.

For the risk assessment component of the CMRA methodology, the computer program FRANCO (Onishi et al. 1979), provides a statistical summarization of time-varying contaminant concentrations predicted by the surface water models discussed for the risk assessment. FRANCO determines the frequency and persistency of specified contaminant concentrations in receiving

waters. Risk is determined by the frequency of occurrence of an event (calculated by FRANCO) and its consequential effects. The consequential effects in the CMRA methodology are expressed in terms of lethality and sublethality by using a median lethal concentration (LC50) and chronic limits; for example, the maximum acceptable toxicant concentration (MATC). By selecting specific concentration-duration levels to match LC50s and MATC values to define a concentration-duration curve (such as a curve connecting LC50 values), FRANCO provides the number of times, duration, and frequency that acute kill, chronic kill, potential acute damage, chronic damages, and no effect occur at given locations in surface water (fig. 10).

The CMRA methodology was applied to various sites including the Four Mile/Wolf Creek watershed in Iowa and the Yazoo River basin in Mississippi to estimate overland/instream migration and associated aquatic impacts of pesticides (Onishi et al. 1985).

The Multimedia Contaminant Environmental Exposure Assessment (MCEA) methodology was developed to realistically assess exposures of air, soil, groundwater, and surface water to chemicals released from stacks, land disposal, disposal ponds, and direct liquid discharges to surface water. The MCEA methodology, which consists of a series of physics-based pathway models to simulate dominant mechanics of chemical migration fate, includes:

(1) the RAPT (McNaughton and Powell 1981), ANDEP (Vaughan et al. 1975), STRAM (Hales et al. 1977), and ISC (Bowers et al. 1979) models to simulate atmospheric transport and chemical deposition and resuspension,
(2) the ARM model (Crawford and Donigian 1973) for the overland pathway,
(3) the UNSAT (Gupta et al. 1978), VTT (Reisenauer 1979), and MMT models (Ahlstrom et al. 1977) for saturated and unsaturated groundwater pathways, and
(4) the TODAM (Onishi et al. 1982b), SERATRA (Onishi et al. 1980), and EXAMS (Smith et al. 1977) models with appropriate hydrodynamic models for the surface water pathway.

Although these specific pathway models were selected for the MCEA methodology, they can be replaced or supplemented when more appropriate models (including geochemical models) become available. The MCEA methodology acts as a framework for multimedia pathway modeling.

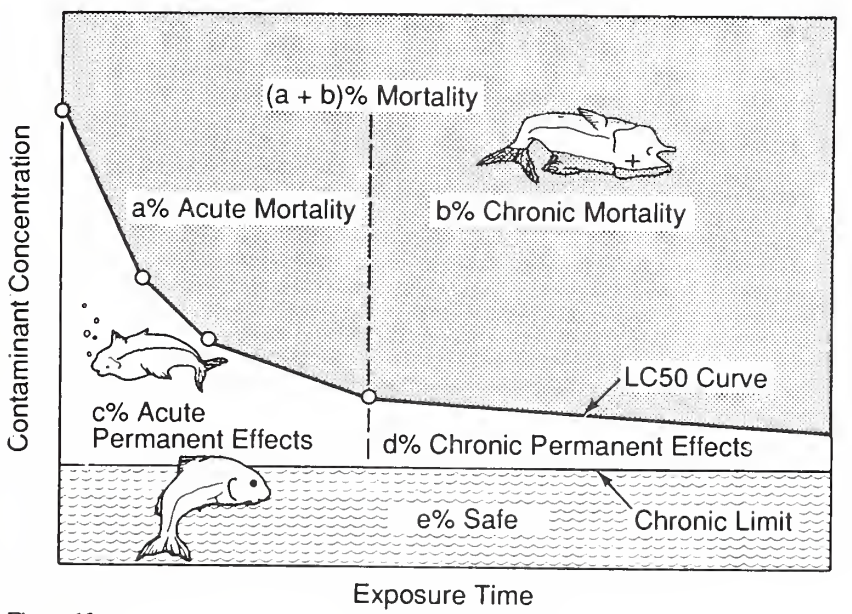


Figure 10.
Risk assessment of the CMRA methodology (Onishi et al. 1985).

For each of the four release modes discussed, necessary procedures were developed to couple models for the primary and secondary pathways of chemicals. On the basis of formulations and past applications, most pathway models selected for the MCEA methodology are approximately equally sophisticated in the ability to simulate dominant dynamic processes of chemical migration and fate in the environment. However, time frames and land surface areas for these models differ considerably. Because it was applied to a hypothetical coal-fired power plant by using only one pathway, the usefulness and practicability of the MCEA methodology must be further tested by applying all components of the methodology.

SPECIAL CONSIDERATIONS AND DATA NEEDS OF MULTIMEDIA MODELING

Major difficulties of properly simulating multimedia transport and fate processes arise from difficulties of treating mass and energy transfer across various pathways, and from different time and space scales for contaminant transport and fate processes in various pathways.

Regardless of the level of spatial description of transport within each environmental medium, multimedia models are driven by the transport of contaminant mass across media boundaries. Thus, an accurate description of the different intermedia transport routes (table 1) (Cohen 1987a) is essential. Intermedia transport parameters are affected by meteorological media properties as well as the physicochemical properties of the pollutant. For example, Ryan and Cohen (1986) showed that the application of a simple multimedia model with accurate predictions of intermedia transport parameters yielded an accurate description of the multimedia distribution of benzo(a)pyrene in the southeast Ohio region. To simplify the incorporation of transport parameter predictions, Cohen and coworkers (Cohen 1981, 1987b; Cohen and Ryan 1985; Ryan and Cohen 1986; Mayer 1988) developed a series of auxiliary-transport-parameters modules that were built into the fully coupled, integrated MCM and SMCM models. The various prediction methods and their applicability and accuracy have been discussed in detail by Cohen (1987a) and Mayer (1988).

A number of areas of intermedia transport require further work to improve our understanding of the mechanisms of intermedia transport and thus the accuracy of current prediction methods. For example, little is known about the wet scavenging of reactive organics, and even the prediction of dry deposition for organic species on rough terrains involves a large degree of uncertainty. Introduction of aerosols into the atmosphere from the oceans and the resuspension of aerosols are also often ignored in multimedia models because of lack of proven prediction methods. Other areas of uncertainty as important as these few examples have been discussed in detail by Cohen (1987c), who provided a list of suggestions for needed research in intermedia transport.

For the partially coupled, integrated model and especially for composite multimedia models, the following major problem areas are associated with coupling one submodel/code with another:

(1) data base
(2) time resolution difference
(3) unit conversion
(4) space resolution differences.

The data base problem is one of ensuring that input data that are used as input to more than one model are taken from the same source whenever possible. In addition, if these input data must be used at different times or spatial scales in the different models, the appropriate time or spatial averaging methods must be used to maintain common material or energy balance between the various codes.

Table 1.
Summary of major intermedia transport routes (Cohen 1987a).

A.	Transport from the Atmosphere to Land and Surface Water
1.	Dry deposition of particulate and gaseous pollutants
2.	Precipitation scavenging of gases and aerosols
3.	Adsorption onto particulate matter and subsequent dry and wet deposition
B.	Transport from Surface Water to Atmosphere, Sediment, and Organisms
1.	Volatilization
2.	Aerosol formation at the air/water interface
3.	Sorption by sediment and suspended solids
4.	Sedimentation and resuspension of solids
5.	Uptake and release by biota
C.	Transport from Soil to Surface Water, Sediment, Atmosphere, or Biota
1.	Dissolution in rainwater
2.	Adsorption on soil particles and transport by runoff or wind erosion
3.	Volatilization from soil and vegetation
4.	Leaching into groundwater
5.	Resuspension of contaminated soil particles by wind
6.	Uptake by microorganisms, plants, and animals
D.	Transport from Groundwater to Surface Water or Biota
1.	Seepage to surface water
2.	Root uptake by plants

The various models are chosen so that one model's output acts as a source input or boundary condition to the next model. These models generally have different time scales and spatial resolutions. For example, surface water models of river systems in many times discretize modeling area into one-dimensional space and are very dynamic in time (often requiring time steps in minutes) while models for regional groundwater systems generally require a modeling area to be at least two dimensions (and time steps that may be as long as a month). In terms of coupling, the problem involves ensuring mass and energy continuity among the various models. Coupling between the various models of the multimedia system can be accomplished by small data manipulation codes, which will take the output from the donor models and perform the proper time and spatial averaging or interpolation along with the appropriate unit conversion in order to prepare the input stream for the next "receiving" model.

In the case of the MCEA methodology, coupling processes from one pathway to others are determined and a series of small data manipulation codes were developed to handle intermedia transfer. Table 2, as an example, shows coupling considerations for surface water pathway models (Onishi et al. 1982b). In contrast, the fully coupled, integrated multimedia models do not require this procedure, because the simultaneous solution of the mass transport equation, subject to the appropriate intermedia boundary conditions, ensures that the mean balance is always satisfied.

Table 2.
Coupling considerations for SERATRA and TODAM (surface water pathway)
(from Onishi et al. 1982b).

Coupling Input From Other Models	Spatial Discretization Requirements	Time Discretization Requirements	Unit Conversion Requirements
RAPT (Air)	Conversion from NMC to cartesian grid system	Linear interpolation for time scale differences	Convert deposition ($\mu g/m^2$) to loading rate (kg/s)
STRAM (Air)	Conversion from NMC to cartesian grid system	Linear interpolation for time scale differences	Convert deposition ($\mu g/m^2$) to loading rate (kg/s)
ANDEP	Conversion from polar to cartesian grid system	Linear interpolation for time scale differences	Convert deposition ($\mu g/m^2$) to loading rate (kg/s)
ISC (Air)	Conversion from polar cartesian grid system	Linear interpolation for time scale differences	Convert deposition ($\mu g/m^2$) to loading rate (kg/s)
ARM (Overload)	Conversion from polar to cartesian grid system	Linear interpolation for time scale differences	Loading rate (kg/min) and loading rate (kg/s) distribute bulk sediment and bulk particulate contaminant into size fractions
MMT (Groundwater)	Conversion from polar to cartesian grid system	Linear interpolation for time scale differences	No conversion necessary

SERATRA and TODAM Output	1. Sediment concentrations in water column (kg/m^3) 2. Dissolved contaminant concentration in water column (kg/m^3) 3. Particulate contaminant concentrations in water column (kg/m^3) 4. Bed elevation change from sediment erosion and deposition (m) 5. Sediment size distributions in river bed (%) 6. Particulate contaminant concentrations in river bed (kg/kg)
SERATRA and TODAM Output	1. Spatial and temporal distributions of contaminant concentrations (kg/m^3) 2. Spatial and temporal distributions of contaminant sorbed to sediment (kg/kg)

THE MULTIMEDIA MODELING APPLICATIONS

To illustrate multimedia modeling applications, two models are discussed here. the fully coupled, integrated multimedia model, represented by the SMCM model, and the composite multimedia model, represented by the CMRA methodology.

MCM and SMCM Model Applications

Mayer (1988) applied the SMCM model to Santa Clara Valley in northern California to determine distributions of the gaseous organic chemical, benzene. The Santa Clara Valley stretches across approximately 3540 square kilometers. As was shown in figure 5, the SMCM model has six compartments representing air, soil, surface water, suspended sediment, surface water bottom sediment, and aquatic biota (fish). Tables 3 and 4 show data for each compartment and the physicochemical characteristics of benzene (Mayer 1988). The model calculates intermedia transfer parameters (table 5). With this information, the model calculates concentrations of benzene in Santa Clara Valley under conditions of no rainfall and rainfall randomly generated. Figure 11, showing the no-rainfall case, indicates that air and water compartments reach their steady states rapidly, while soil reaches the steady state most slowly because of smaller diffusion. The effect of rainfall in Santa Clara Valley was found to have negligible effect on contaminant transport. Study results also indicate that most of the benzene remained in the air compartment, and its concentration, 10.7×10^{-8} mol/m^3 (7.1 ppb), matched fairly well with monitored values of 7.7×10^{-8} mol/m^3 (5.1 ppb).

Although the model does not solve spatial distributions in four of the six pathways, this model application nevertheless demonstrates the usefulness of the model to predict the response of each pathways to the benzene originally released to air. A similar simulation for TCE distribution in the Santa Clara region (Mayer 1988) and in the La Jolla area (Cohen and Ryan 1985) also demonstrated the usefulness of the fully coupled, integrated models for screening analysis. The simple MCM-type models can also be applied to particle-bound pollutants, in the case of the MCM model used to predict b(a)p distribution in the southeast Ohio region (Ryan and Cohen 1986).

Table 3.
Compartmental data for Santa Clara Valley (from Mayer 1988).

Physical Data	Air	Water	Soil	Sediment	Fish	Suspended Solids
Volume (m^3)	1.4×10^{12}	1.9×10^{8}	2.8×10^{10}	1.9×10^{7}	96	960
Flow Rates (m^3/hr)	3×10^{9}	--	--	--	--	--
Densities (g/cm^3)[1]	--	1.0	1.5[1]	1.5	--	1.5

[1]Density of soil and sediment solid phases.

Table 4.
Physicochemical data for benzene (from Mayer 1988).

Parameter	Benzene
Source in air	975 mol/hr
Degradation in air	43×10^{-4} mol/hr^{-1}
Degradation in water	48×10^{-4} mol/hr^{-1}
Air background concentration	1.5×10^{8} mol/m^3
Molecular weight	78.12 g/mol
Solubility	0.023 mol/l
Henry's law coefficient	555 Pa•m^3/mol
Normal boiling temperature	353.25 K
Molal volume	89 cm^3/g•mol

Table 5.
Parameters calculated by the SMCM model (from Mayer 1988).

	Benzene
Partition Coefficients	
Water/Air	4.2
Soil/Air	6.7
Biota/Water	4.6
Suspended Solids/Water	2.5
Sediment/Water	1.8
Mass Transfer Coefficients (m/hr)	
Air/Water	0.315
Water/Air	0.074
Air/Soil	0.072
Soil/Air	0.011
Biota/Water	9.83
Suspended Solids/Water	0.266
Sediment/Water	0.0015
Diffusion Coefficients (m^2/hr)	
Dry Soil	9.9×10^{-4}
Wet Soil	9×10^{-6}
Sediment	4.0×10^{-7}

CMRA Methodology Application

The CMRA methodology was applied to the Yazoo River basin, Mississippi, to assess the potential risks to fish resulting from assumed applications of a persistent chlorinated hydrocarbon insecticide, toxaphene, to farmland in the basin (Onishi et al. 1985).

This study assisted the U.S. Environmental Protection Agency (EPA) in deciding whether the use of toxaphene in the United States should be restricted or banned. Specifically, toxaphene migration and fate were evaluated in the Coldwater, Tallahatchie, Yazoo, and Big Sunflower rivers (fig. 12), and the potential impacts of toxaphene on four fish [largemouth bass (Micropterus salmonides), bluegill sunfish (Lepomis macrochirus), fat-head minnow (Pimephales promelas), and channel catfish (Ictalurus punctatus)] were assessed in these rivers for March 1971 to December 1975.

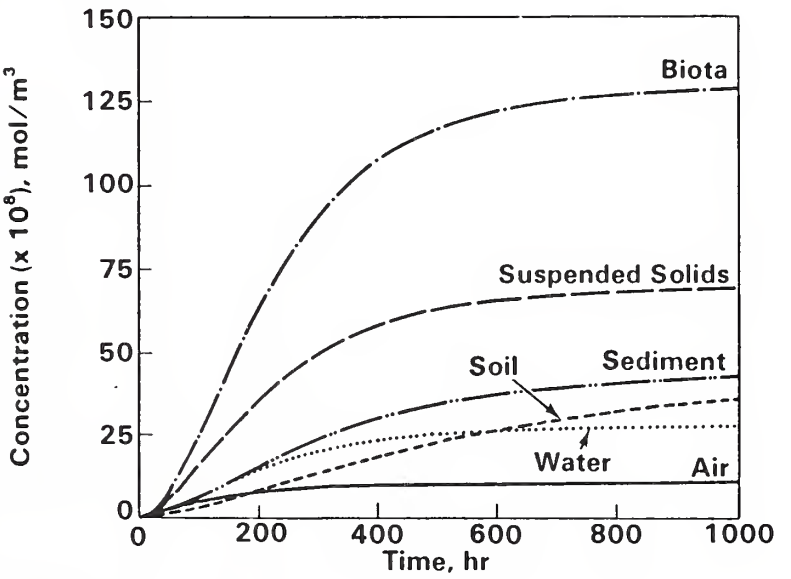


Figure 11.
Distribution of benzene in Santa Clara Valley: no-rainfall case (SMCM model) (Mayer 1988).

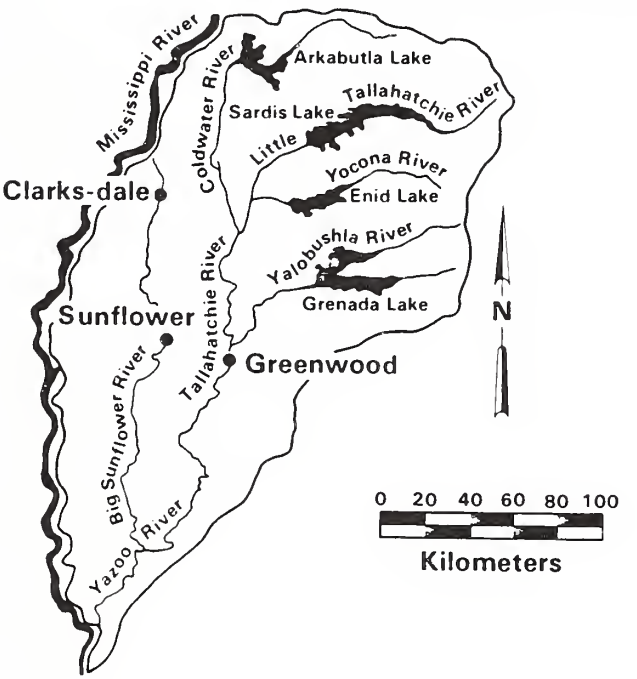


Figure 12.
Yazoo River Basin, Mississippi.

Toxaphene in these rivers is mostly absorbed by fine sediment (e.g., silt, clay, and organic matter). Fine sediment usually is transported through the river system without deposition except in very slow-moving areas. Thus, the Mixed Tank Model, which does not allow sediment deposition or resuspension, was applied to these relatively small and fast-moving rivers instead of more detailed (sediment and contaminant transport) codes, such as TODAM and SERATRA of the CMRA methodology. The suspended sediment adsorbs dissolved pesticide, resulting in a smaller dissolved concentration. Thus, for a given total pesticide concentration, the higher the suspended sediment concentration, the lower the dissolved concentration. Sediment and toxaphene loadings from farmland to each of the reaches of these rivers were calculated by the EPA Environmental Research Laboratory (Athens, Georgia) using the ARM model.

Computed dissolved (mg/l) and particulate (sediment-sorbed) (mg/g) toxaphene concentrations near the mouth of the Yazoo River are shown in figures 13 and 14, respectively, showing large temporal variations of the concentrations. Based on the LC50 and MATC values obtained from laboratory toxicological tests, the FRANCO program of the CMRA methodology statistically summarized the predicted dissolved toxaphene concentrations to provide risk assessment for these fish. Note that the risk assessment was somewhat hampered by a literal interpretation of the laboratory toxicology to the actual field conditions. Figure 15 shows probabilistic risk assessment results for young largemouth bass (most affected). As indicated in figure 15, approximately 15% of the time during the 5-year simulation period, the dissolved toxaphene concentrations were above the LC50 curve for young largemouth bass (i.e., for approximately 15% of the time, more than 50% of young largemouth bass would be killed). FRANCO indicated that this situation would occur 18 times, totaling 60 days, during the simulation. For almost 60% of the time the dissolved toxaphene fell between the LC50 curve and the MATC level of 0.072 μg/l, indicating the expectation of some lethal and sublethal damages to largemouth bass. During that time, dissolved toxaphene levels remained mostly between the LC50 curve and the MATC value for longer than 4 days (96 hr). This reveals that most of the impact would occur as chronic damage. However, for 27% of the simulation time, dissolved toxaphene concentration was less than the MATC value, during which time largemouth bass would be safe from measurable effects. Other fish (bluegill sunfish, fathead minnow, and channel catfish) are more resistant to toxaphene exposure.

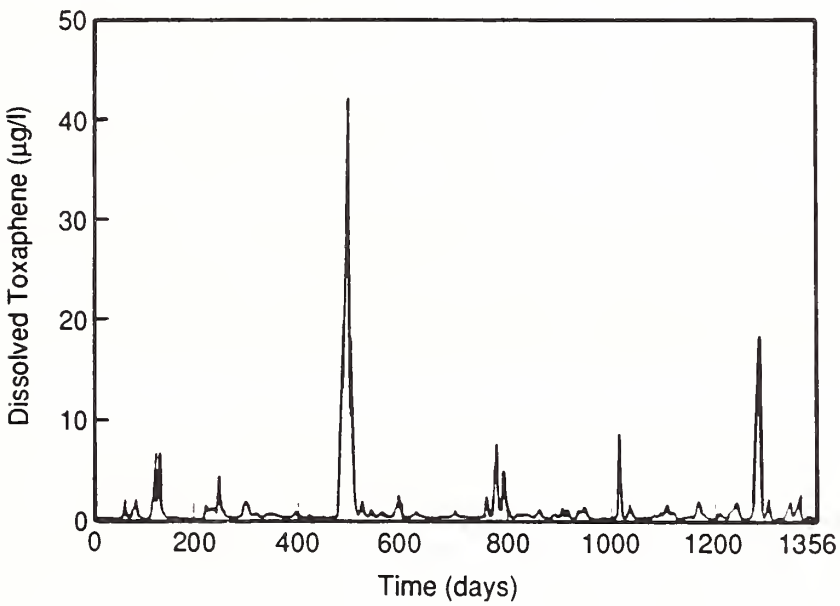


Figure 13.
Predicted dissolved toxaphene concentration near the mouth of the Yazoo River.

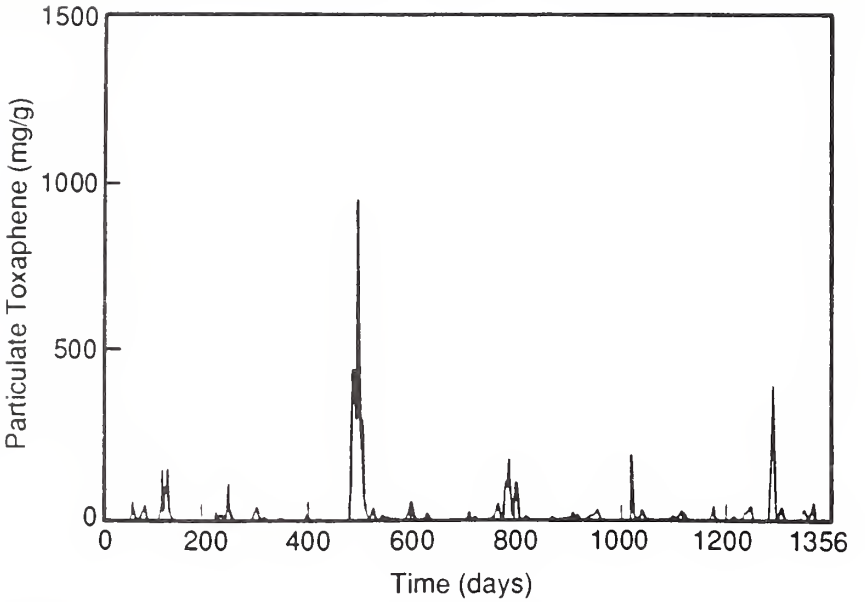


Figure 14.
Predicted particulate toxaphene concentrations (mg/g) near the mouth of the Yazoo River.

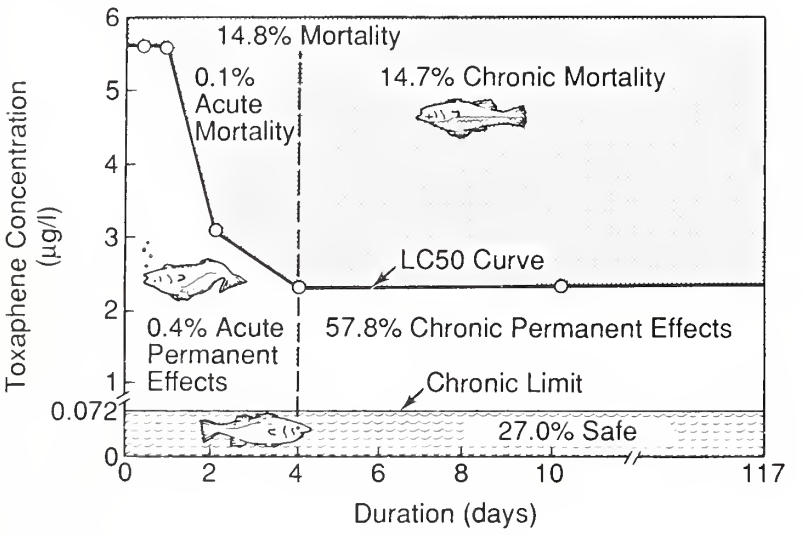


Figure 15.
Probabilistic risk assessment of young largemouth bass.

This assessment by the CMRA methodology revealed both acute and chronic damage to these fish. Partially because of these study results, the U.S. government banned the use of toxaphene in the United States. This study thus reveals that a composite multimedia model is useful to assess potential problematic cases.

CONCLUSIONS AND FUTURE RESEARCH RECOMMENDATIONS

Toxic chemicals released to the environment undergo complex interactions among various transport and fate mechanisms and multiple pathways (see figs. 1 and 2). Consequently, multimedia models are very useful tools for integrating transport and fate mechanisms in a single framework to assess contaminant distribution in the environment. These models must adequately

represent spatial and temporal variability of transport, fate, and intermedia transfer processes. They can be divided into three groups: fully coupled, integrated multimedia models; partially coupled, integrated multimedia models; and composite multimedia models. The first group emphasizes the intermedia transfer of chemicals and potential temporal variations. In contrast, the third group emphasizes transport processes in each environmental pathway and spatial variations of chemicals. The second group is intermediate in concept. Each of these groups of models has specific objectives, and none covers all conditions and objectives.

Many restrictions of multimedia modeling arise from the difficulties of properly evaluating intermedia transfer of mass and energy and of coupling different temporal and spatial scales of contaminant transport and fate processes occurring in various environmental pathways. Another difficulty is that of simulating, with reasonable accuracy, the contaminant transport and fate in the environment during time frames spanning several weeks (e.g., Chernobyl) to several decades. Future studies must address these problems. Finally, to properly test the ability of multimedia transport models to predict pollutant distribution under field conditions, appropriate multimedia field monitoring data must be acquired.

REFERENCES

Ahlstrom, S.W., H.P. Foote, R.C. Arnett, and others. 1977. Multicomponent Mass Transport Model: Theory and Numerical Implementation (Discrete-Parcel-Random-Walk Version). BNWL-2127, Pacific Northwest Laboratory, Richland, Washington.

Bolten, J.G., P.F. Morrison, and K.A. Soloman. 1983. Risk-Cost Assessment Methodology for Toxic Pollutants from Coal-Fired Power Plants. WD-1539/EPRI RP1826-5, Electric Power Research Institute, Palo Alto, California.

Bowers, J.F., J.R. Bjorkland, and C.W. Cheroy. 1979. Industrial Source Complex (ISC) Dispersion Models User's Guide. EPA-450/4-79-030, U.S. Environmental Protection Agency, Washington, D.C.

Cohen, Y. 1981. Systems Approaches: Compartmental Modeling. NCITR Report CR-807864-01-1, National Center for Intermedia Transport Research, University of California at Los Angeles, California.

Cohen Y. 1987a. Intermedia Transport Modeling in Multimedia Systems. National Center for Intermedia Transport Research, University of California at Los Angeles, California.

Cohen, Y. 1987b. Pollutants in a Multimedia Environment, ed. Y. Cohen. Plenum Press, New York.

Cohen, Y. 1987c. In Geochemical and Hydrologic Processes and Their Protection - The Agenda for Long-Term Research and Development, eds. S. Draggan, J.J. Cohrssen, and R.E. Morrison. Praeger, New York.

Cohen, Y., and P.A. Ryan. 1985. Multimedia modeling of environmental transport: trichloroethylene test case. Environ. Sci. Technol. 19:412.

Crawford, N.H., and A.S. Donigian, Jr. 1973. Pesticide Transport and Runoff Model for Agricultural Lands. EPA 660/2-74-013, Office of Research and Development, U.S. Environmental Protection Agency, Washington, D.C.

Donigian, A.S., Jr., and N.H. Crawford. 1976. Modeling Pesticides and Nutrients on Agricultural Lands. EPA-600/2-76-043, U.S. Environmental Protection Agency, Athens, Georgia.

Donigian, A.S., Jr., J.C. Imhoff, B.R. Bicknell, and others. 1983. Application of Hydrologic Simulation Program FORTRAN (HSPF) in Iowa Agricultural Watersheds. EPA-600/S3-83-069, U.S. Environmental Protection Agency, Athens, Georgia.

EPA (U.S. Environmental Protection Agency). 1986. Hazardous Waste Management System Land Disposal Restrictions; Proposed Rule. Part III. U.S. Environmental Protection Agency, 40 CFR (260) Fed. Reg. 1601-1766 (January 14, 1986).

Fletcher, J.F., and W.L. Dodson. 1971. Hermes - A Digital Computer Code for Estimating Regional Radiological Effects from the Nuclear Power Industry. USAEC Rep. HEDL-TM#-71-1968, U.S. Atomic Energy Commission, Washington, D.C.

Gupta, S.K., K.K. Tanji, D.R. Nielsen, and others. 1978. Field Simulation of Soil-Water Movement with Crop Water Extraction. Water Science and Engineering Paper No. 4013, Department of Land, Air and Water Resources, University of California, Davis, California.

Hales, J.M., D.C. Powell, and T.D. Fox. 1977. STRAM - An Air Pollution Model Incorporating Nonlinear Chemistry, Variable Trajectories and Plume Segment Diffusion. EPA-450-13-77-012, U.S. Environmental Protection Agency, Athens, Georgia.

Johanson, R.C., J.C. Imhoff, and H.H. Davis, Jr. 1980. User's Manual for Hydrologic Simulation Program - FORTRAN (HSPF). EPA-600/9-80-015, U.S. Environmental Protection Agency, Athens, Georgia.

Knisel, W.G. 1980. CREAMS: A Field-Scale Model for Chemicals, Runoff, and Erosion from Agricultural Management Systems. Conserv. Res. Rep. No. 26, U.S. Department of Agriculture, Washington, D.C.

Mackay D. 1987. Pollutants in a Multimedia Environment. ed. Y. Cohen. Plenum Press, New York.

Mackay, D., and S. Paterson. 1982. Fugacity revisited. The Fugacity Approach to Environmental Transport. Environ. Sci. Technol. 16:654A-660A.

Mayer, G.J. 1988. The Spacial Multimedia Compartmental Model (SMCM): A User-Friendly, Screening Level Pollutant Distribution Model. Master's Thesis, Chemical Engineering Department, University of California at Los Angeles, California.

McKone, T.E. 1985. The Use of Environmental Health-Risk Analysis for Managing Toxic Substances. Presented at the 78th Meeting of the Air Pollution Control Association, Detroit, Michigan, June 16-21, 1985 (Lawrence Livermore National Laboratory Report UCRL-92329).

McKone, T.E., and W.E. Kastenburg. 1985. Application of Multimedia Pollutant Transport Models to Risk Analysis. RARE 85-169.

McNaughton, D.J., and D.C. Powell. 1981. RAPT - The Pacific Northwest Laboratory Regional Air Pollutant Transport Model: A Guide. PNL-3390, Pacific Northwest Laboratory, Richland, Washington.

Onishi, Y. 1981. Sediment-Contaminant Transport Model. ASCE J. Hydraul. Div. 107:1089-1107.

Onishi, Y., and D.S. Trent. 1985. Three-Dimensional Simulation of Flow, Salinity, Sediment, and Radionuclide Movements in the Hudson River Estuary. In Proceedings of the ASCE Hydraulics Division Specialty Conference, Lake Buena Vista, Florida, August 12-17, 1985, pp. 1095-1100.

Onishi, Y., and S.E. Wise. 1979. Finite Element Model for Sediment and Toxic Contaminant Transport in Streams. In Proceedings of Conservation and Utilization of Water and Energy Resources, pp. 144-150. Hydraulics and Energy Divisions, ASCE, San Francisco, California.

Onishi, Y., D.L. Schreiber, and R.B. Codell. 1980. Mathematical Simulation of Sediment and Radionuclide Transport in the Clinch River, Tennessee. In Processes Involving Contaminants and Sediment, ed. R.A. Baker, pp. 393-406, Chap. 18. Ann Arbor Science, Ann Arbor, Michigan.

Onishi, Y., S.M. Brown, A.R. Olsen, and others. 1979. Assessment Methodology for Overland and Instream Migration and Risk Assessment of Pesticides. Battelle, Pacific Northwest Laboratories, Richland, Washington.

Onishi, Y., S.M. Brown, A.R. Olsen, and M.A. Parkhurst. 1981a. Chemical Migration and Assessment Methodology. In Proceedings of the Conference on Environmental Engineering, American Society of Civil Engineers, Atlanta, Georgia.

Onishi, Y., R.J. Serne, E.M. Arnold, C.E. Cowan, and F.L. Thompson. 1981b. Critical Review: Radionuclide Transport, Sediment Transport, and Water Quality Mathematical Modeling, and Radionuclide Adsorption/Desorption Mechanisms. NUREG/CR-1322/PNL-2901, Pacific Northwest Laboratory, Richland, Washington.

Onishi, Y., G. Whelan, and R.L. Skaggs. 1982a. Development of a Multimedia Radionuclide Exposure Model for Low-Level Management. PNL-3370, Pacific Northwest Laboratory, Richland, Washington.

Onishi Y., S.B. Yabusaki, C.R. Cole, W.E. Davis, and G. Whelan. 1982b. Multimedia Contaminant Environmental Exposure Assessment (MCEA) Methodology for Coal-Fired Power Plants. vols. I and II. Battelle, Pacific Northwest Laboratories, Richland, Washington.

Onishi, Y., A.R. Olsen, M.A. Parkhurst, and G. Whelan. 1985. Computer-Based Environmental Exposure and Risk Assessment Methodology for Hazardous Materials. J. Hazard. Mater. 10:389-417.

Patterson, M.R. 1986. Unified Transport Model for Organics. In Proceedings of the First Workshop on Pollutant Transport and Accumulation in a Multimedia Environment. January 22-24, 1986, Los Angeles, California.

Reisenauer, A.E. 1979. Variable Thickness Transient Ground-Water Flow Model, vol. 1. PNL-3160-1, Pacific Northwest Laboratory, Richland, Washington.

Ryan, P.A. 1984. Multimedia Modeling of Environmental Transport. M.S. Thesis, Chemical Engineering Department, University of California at Los Angeles, California.

Ryan, P.A., and Y. Cohen. 1986. Multimedia Benzo[a]pyrene Test Case. Chemosphere 15:21-47.

Smith, J.H., W.R. Mabey, N. Bohonos, and others. 1977. Environmental Pathways of Selected Chemicals in Freshwater Systems. Part I. Background and Experimental Procedures. EPA-600/7-77-113. Prepared for U.S. Environmental Protection Agency by SRI International, Menlo Park, California.

Soldat, J.K., N.M. Robinson, and D.A. Baker. 1974. Models and Computer Codes for Evaluating Radiation Doses. BNWL-1754, Pacific Northwest Laboratory, Richland, Washington.

Tucker, W.A., A.Q. Eschenroeder, and G.C. Magill. 1984. Air Land Water Analysis System (ALWAS): A Multi-Media Model for Toxic Substances. EPA-600/S3-84-052/NTIS PB 84-171 743, U.S. Environmental Protection Agency, Athens, Georgia.

Vaughan, B.E., and others. 1975. Review of Potential Impact on Health and Environmental Quality from Metals Entering the Environment as a Result of Coal Utilization. Battelle Energy Program Report, Battelle-Northwest, Richland, Washington.

Whelan, G., and Y. Onishi. 1983. In-Stream Contaminant Interaction and Transport. In Proceedings of the Tenth International Symposium on Urban Hydrology, Hydraulics, and Sediment Control, July 25-28, 1983, University of Kentucky, Lexington, Kentucky.

Whelan, G., and M.A. Parkhurst. 1983. Simulation of the Migration, Fate and Effects of Diazinon In-Stream. In Proceedings of the D.B. Simons Symposium on Erosion and Sedimentation, July 27-29, 1983, Colorado State University, Fort Collins, Colorado.

Whelan, G., D.L. Strenge, J.G. Droppo, Jr., and others. 1987. The Remedial Action Priority System (RAPS). Mathematical Formulation. DOE/RL/87-09/ PNL-6200, Pacific Northwest Laboratory, Richland, Washington.

DISCUSSION OF PAPERS PRESENTED IN TECHNICAL SESSION 5: INTEGRATED WATER QUALITY MODELS FROM THE MODELER'S AND USER'S PERSPECTIVES

L.D. James[1], Presiding
D.K. Stevens[2], Recorder

PAPERS DISCUSSED

Environmental Isotope Tracer Studies of Catchment Processes: Tools for Testing Integrated Water Quality Models by M.G. Sklash, I.D. Moore and G.J. Burch

Multimedia Modeling of Toxic Chemicals by Y. Onishi, L. Shuyler and Y. Cohen

SPECIFIC QUESTIONS AND COMMENTS

Question: (Reese) Could the scatter in the ^{127}Ce and ^{7}Be data be due to variability in the labeled and unlabeled soils?

Response: (G. Burch, Bureau of Rural Resources, Australia) New data are forthcoming that may help shed some light on this point. In one study, nearly all the material analyzed was from a gully and fairly uniform in composition. The use of Sr isotopes as tracers may provide more sensitive indicators to further separate the two types of flow.

Question: (T. Richard, Cornell University, Ithaca, New York) In the New Zealand studies cited, 20-40% of the flow was new water; where did the old water, on the steep slopes of the study catchment, come from?

Response: (M. Sklash, Department of Geology, University of Windsor, Canada) Sufficient water resided in the soil in valley bottoms to account for the difference.

Question: (T. Richard) Does the water run out of the soil into macropores?

Response: (M. Sklash) A study in Scotland demonstrated that it is simply forced out through the soil matrix by the new water.

Question: (J. Troiano, California Department of Food and Agriculture, Sacramento, California) Was there any data concerning toxaphene cited in your talk?

Response: (Y. Onishi, Batelle Pacific NW Labs, Richland, Washington) The results shown were simulated.

Question: (N. Jarvis, SW University Agricultural Science, Sweden) Are you certain that macropore flow wasn't an important contributor to the total flow? Vertical macropores could be important and if so, old water would be forced out. Could shallow water tables be important contributors?

Response: (M. Sklash) They are, although a study in California suggested the opposite.

[1]L.D. James, Director and Professor, Utah Water Research Laboratory, Utah State University, Logan, Utah.
[2]D.K. Stevens, Department of Civil and Environmental Engineering, Utah State University, Logan, Utah.

Comment: (G. Burch) Based on Australian studies, the response of the shallow water table was very dynamic during rainfall events. Modeling of the response showed that it was dominated by the watershed.

Question: (D. Woolhiser, USDA-ARS, Tucson, Arizona) How was the rainfall simulated in the overland flow model?

Response: (Y. Onishi) In the ARM model, rainfall was uniformly distributed over time and the study area.

Comment: (I. Moore, Department of Engineering, University of Minnesota, St. Paul, Minnesota) In the context of modeling the erosion process, the available tools were limited - one works with average prediction rather than attempting to model spatial variability.

Question: (D. James, Utah Water Research Laboratory, Utah State University, Logan, Utah) Very often, groundwater flow intersects hazardous water disposal sites. Would the radiotracer technique be useful in this application?

Response: (M. Sklash) The isotopes are currently used to test models. If the old water in the landfill has a different isotopic content than the base flow, it would be possible to trace where the contaminant plume was going. Lithium and ^{14}C tracers would be useful.

Comment: (Y. Onishi) The multimedia models are also useful in this regard. The RAPS model has been used for toxics to screen and rank hazardous waste disposal sites for remediation. They can also be used to predict contamination movement in the environment and perform risk assessments.

Comment: (I. Moore) In the Murray Basin, Australia, much of the increased salinity was due to land clearing and resultant increase in recharge in that large basin. Tritium and ^{36}Cl have been used as tracers to better estimate small changes in recharge.

Question: (H. Roaza, Florida Department of Environmental Regulation, Tallahassee, Florida) Has ^{18}O been used to date the water?

Response: (M. Sklash) The reported work was only in prediction of the mixing of old and new water.

Comment: (H. Roaza) ^{18}O measurements had been used to study the mixing of marine and fresh waters and the measurements were highly temperature-sensitive.

Comment: (M. Sklash) The temperature of measurement is important but in our work the errors were less than 10%.

Comment: (G. Burch) In estimating recharge in arid zones, measurement of temperature profiles in the soil was essential.

Question: (A. Patwardhan, Minnesota Pollution Control Agency, St. Paul, Minnesota) What about the difficulty in transferring measurements between media?

Response: (Y. Onishi) The interchange of contaminants between media in the model was in the form of specifying fluxes as sink and source terms in compartmental models. In fully-integrated models the fluxes are managed by mass transfer rate equations with a driving force equal to the product of the concentration gradient and a mass transfer coefficient. Values of the mass transfer coefficients are usually estimated within the fully-integrated models.

EFFECTS OF SPATIAL AND TEMPORAL VARIABILITY ON WATER QUALITY MODEL DEVELOPMENT

D.A. Woolhiser[1], W.A. Jury[2], and D.R. Nielsen[3]

ABSTRACT

The deterministic partial differential equations describing transport of agricultural chemicals in surface runoff, in the vadose zone and in the groundwater zone are briefly reviewed. Significant spatial and temporal variations of input and parameter fields are identified and modeling techniques to incorporate these variations at various scales are discussed. It is concluded that explicit procedures must be devised to include small scale spatial variability into agricultural chemical transport models without excessive computational effort.

INTRODUCTION

Development of hydrologic transport or water quality models began about twenty years ago and the pace of research and application of these models has accelerated appreciably as society has become more aware of the serious environmental problems associated with pollutants in water supplies. The principal goal of water quality models is to understand and predict the movement of various chemicals in surface and subsurface flow to provide information for policy decisions or remedial activities. The questions that water quality models should answer are: how much of the chemical will reach surface waters or ground waters; where will it go; when will it arrive; and, in remedial activities, when will it no longer be a problem?

These are difficult questions to answer under the best of circumstances but when we consider the complexities introduced by the stochastic nature of the inputs and the temporal and spatial variations of the input and parameter fields it becomes obvious that considerable simplification is required to develop chemical transport models. The challenge to researchers is to identify the important physical, chemical and biological processes involved and provide a mathematical description of these processes and their linkages at an appropriate scale. The most important temporal and spatial variations in input and parameter fields must also be identified and these variations must be explicitly accounted for in model construction.

For purposes of discussion we will partition a catchment into three zones: surface, vadose and groundwater. It is clear that this is an artificial classification, but modelers have historically made such a distinction. Significant spatial and temporal variations of input fields or parameters and appropriate distance or time scales are summarized for each zone in tables 1 and 2. Obviously there may be some debate about the limits chosen, and we have deliberately left the definition of characteristic length or time somewhat vague, however the tables provide some notion of factors to be considered as we go from laboratory to watershed scale.

[1]D.A. Woolhiser, Research Hydraulic Engineer, USDA-ARS, Tucson, AZ.

[2]W.A. Jury, Professor, Department of Soil and Environmental Sciences, University of California, Riverside, CA.

[3]D.R. Nielsen, Professor, Department of Land, Air, and Water Resources, University of California, Davis, CA.

Table 1.
Spatial Variability.

Parameter or input	Characteristic length (meters)
Precipitation	100 -1000
Evaporation	1 -1000
Chemical application	
Mesoscale (local)	100
Microscale (laboratory)	<1
Surface zone:	
Interception	1 - 100
Infiltration parameters	
Mesoscale (local)	10 - 100
Microscale (laboratory)	<1 - 10
Sources and sinks	
Water	1
Chemical	1
Topography	
Mesoscale (local)	10 - 100
Microscale (laboratory)	.01 - 1
Vadose zone:	
Hydraulic conductivity, porosity, sorptivity, etc.	
Mesoscale (local)	10 - 100
Microscale (laboratory)	.1 - 1
Groundwater zone:	
Hydraulic conductivity, porosity, etc.	
Mesoscale (local)	10 - 100
Microscale (laboratory)	.01 - 1

Table 2.
Temporal Variability.

Parameter or input	Temporal scale
Storm occurrences	Hours - Days
Rainfall intensity	Minutes - Hours
Temperature (air)	Hours - Days
Infiltration parameters	Seasonal, abrupt (tillage operations)
Chemical application	abrupt
Tillage operations	abrupt
Soil physical properties	abrupt (tillage), Weeks - Years
Root growth and decay	Days - Weeks - Years

AGRICULTURAL CHEMICAL TRANSPORT BY WATER

Surface Runoff

Factors governing the partitioning of a chemical species between surface runoff and infiltrating water are poorly understood, yet are crucial in predicting pollution of surface or ground waters. Bailey et al. (1974) described the chemical transport processes involved in the pickup of pollutants (pesticides) by overland flow during a rainstorm, and identified the four important transport mechanisms depicted in figure 1:

(1) Diffusion and turbulent transport of a dissolved chemical species by movement of soil water into the overland flow;
(2) Desorption of the chemical species from soil particles into the soil water or directly into overland flow;
(3) Dissolution of solid phase chemical into soil water or overland flow;
(4) Scouring of solid phase chemical or soil particles by hydraulic forces and subsequent transport and moving dissolution or adsorption-desorption.

The relative importance of each of these transport mechanisms is determined by the chemical under consideration, the method of application, soil characteristics, vegetation and the recent hydrologic history. We will limit our discussion to chemical transport during runoff generated by rainfall. As rainfall begins, individual drops are intercepted by vegetation or fall directly on bare soil and are absorbed. As rain continues to fall at variable rates, the soil water content increases and water moves downward under the influence of gravity and capillary tension gradients. If rainfall rates are sufficiently high, portions of the soil surface will become saturated and the soil will no longer transmit water downward at the rainfall rate. The "excess rainfall" accumulates on the surface and when surface tension effects are overcome, it begins to move downhill under the influence of gravity. Under some circumstances, the topsoil may become saturated and water, which was initially stored in the soil profile or has infiltrated upslope during the storm, may "exfiltrate" and become a contribution to overland flow. The physical and chemical processes involved in the transport of a chemical species from the soil into overland flow are very

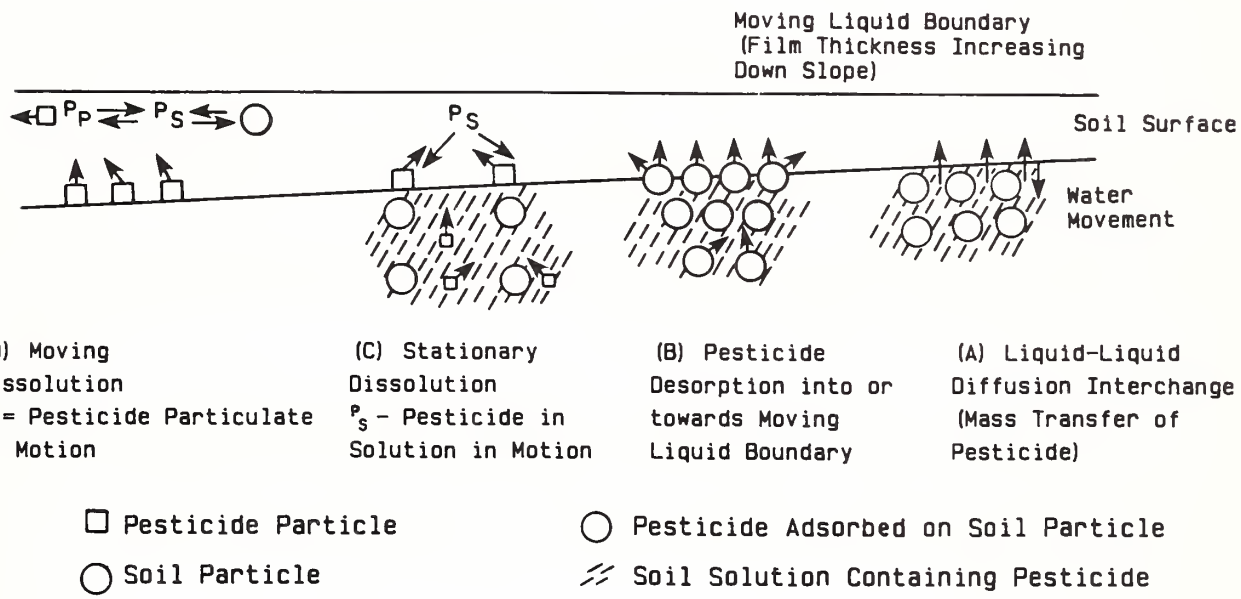


Figure 1.
Modes of pesticide transport into and within the moving liquid boundary during a rainfall event. [Bailey et al. (1974)].

complicated and can be best approached by considering a highly simplified model which requires several strong assumptions. As these assumptions are relaxed, the model will have greater physical and chemical fidelity, purchased at the cost of greater complexity.

A complete description of the process requires mathematical models of the following phenomena:

(1) Two- or three-dimensional unsaturated and saturated flow in a heterogeneous porous medium;
(2) Unsteady, free surface flow over a surface with relatively large roughness elements;
(3) Raindrop impact on the free surface flow and the resulting transient pressure fields developed at the soil/water interface where the pressure and shearing forces developed cause erosion of soil particles and aggregates;
(4) Detachment and transport of soil particles and aggregates by overland and rill flow;
(5) Advective and dispersive movement of chemicals with soil water including chemical and biological decay, adsorption and desorption.

Water movement as surface runoff can be described by the de Saint-Venant equations for wide channels:

$$\frac{\partial h}{\partial t} + \frac{\partial uh}{\partial x} = r(x,t) - f(x,t) \qquad [1]$$

$$\frac{\partial uh}{\partial t} + \frac{\partial (u^2 h)}{\partial x} - gh\left(s_o - \frac{\partial h}{\partial x} - s_f\right) = 0 \qquad [2]$$

where u = cross-sectional average velocity [LT^{-1}],
h = mean flow depth normal to the distance coordinate [L], x,
t = time,
$r(x,t)$ = rainfall rate [LT^{-1}],
$f(x,t)$ = infiltration rate [LT^{-1}],
S_o = bottom slope,
g = gravitational acceleration [LT^{-2}], and
S_f = friction slope.

The friction slope is frequently computed using the Darcy-Weisbach formula:

$$S_f = f_d u^2/8gh \qquad [3]$$

where the friction factor, f_d, depends on the surface roughness, Reynolds number and rainfall intensity.

Equations 1 and 2 are coupled with soil water transport equations at the soil-water interface by requiring continuity of the pressure head and water flux. Prior to ponding, the infiltration rate equals the rainfall rate. After ponding, the pressure head at the surface is equal to the depth of overland flow, h. The coupled overland flow and porous media flow equations can be solved by finite difference equations as demonstrated by Akan and Yen (1981).

Mass transport of a chemical species in overland flow involves four basic mechanisms:

(1) Advection - the time-smoothed mass average flow resulting from the bulk fluid velocity.
(2) Turbulent diffusion - the fluctuating mass flow resulting from the local fluid velocity.
(3) Molecular diffusion - the random migration of individual molecules resulting from their kinetic energy.
(4) Dispersion - introduced by tortuosity of flow paths and temporary storage in depressions.

Turbulent and molecular diffusion are normally insignificant compared with advection and dispersion, and can be neglected. The convection dispersion equation for overland flow can be written as:

$$h\frac{\partial C_A}{\partial t} - \left(\frac{\partial Eh}{\partial x} - uh\right)\frac{\partial C_A}{\partial x} - Eh\frac{\partial^2 C_A}{\partial x^2} + \left(\frac{\partial uh}{\partial x} + \frac{\partial h}{\partial t}\right)C_A = rC_{AR} - fC_A + F \qquad [4]$$

where C_A = local concentration of species A in overland flow [ML^{-3}],
C_{AR} = concentration in rainfall [ML^{-3}],
E = coefficient of dispersion [L^2T^{-1}], and
F = a source or sink term including the mechanisms discussed by Bailey et al. (1974) and identified in figure 1 [$ML^{-2}T^{-1}$].

Sediment transport in overland flow can be described by the convection dispersion equation:

$$h\frac{\partial C_s}{\partial t} - \left(\frac{\partial Eh}{\partial x} - uh\right)\frac{\partial C_s}{\partial x} - E\frac{h\partial^2 C_s}{\partial x^2} + \left(\frac{\partial uh}{\partial x} + \frac{\partial h}{\partial t}\right)C_s = S_R + S_\tau + S_E - S_D \qquad [5]$$

where C_s = concentration of sediment [L^3/L^3],
S_R = rate of sediment entrainment by raindrop impact [LT^{-1}],
S_τ = rate of erosion by hydraulic shear forces [LT^{-1}],
S_E = rate of sediment reentrainment [LT^{-1}], and
S_D = rate of sediment deposition [LT^{-1}].

Transport In The Vadose Zone

The vadose zone, which includes the soil water zone, the intermediate zone and the capillary fringe, may be virtually nonexistent in lowland areas adjacent to bodies of water or may be more than 300 m. thick in arid regions. Within the vadose or unsaturated zone water flows under the influence of matric potential gradients and gravity. The transport equation for water moving through heterogeneous porous media is given by Richards' equation

$$C(\underset{\sim}{r}, \psi)\frac{\partial \psi}{\partial t} = \nabla \cdot \underset{\approx}{K} \cdot \nabla H - S_w \qquad [6]$$

where $C(\underset{\sim}{r}, \psi)$ = water capacity function [L^{-1}],
ψ = matric potential [L],
$\underset{\approx}{K}$ = $\underset{\approx}{K}(\underset{\sim}{r}, \psi)$ is the hydraulic conductivity tensor [LT^{-1}],
$\underset{\sim}{r}$ = (x, y, z) is the position vector,
H = $\psi + z$ is total water potential [L],
z = the vertical coordinate, and
S_w = $S_w(\underset{\sim}{r}, t)$ is a sink term [T^{-1}], generally describing the uptake of water by plant roots.

As a second rank, symmetric tensor $\underset{\approx}{K}$ has six components, each of which is a function of position and matric potential or water content θ.

The corresponding transport equation used for solutes has traditionally been the convection-dispersion model, which for a soluble chemical which adsorbs to soil solids may be written as

$$\frac{\partial}{\partial t}(\theta C_\ell + \rho_b C_s) = \underset{\sim}{\nabla} \cdot \underset{\approx}{D} \cdot \underset{\sim}{\nabla} C_\ell - \underset{\sim}{J_w} \cdot \underset{\sim}{\nabla} C_\ell - S_s \qquad [7]$$

where θ = the water content [L^3/L^3]

C_ℓ[ML^{-3}] and C_s [M/M] = dissolved and adsorbed concentrations, respectively,

$\underset{\sim}{J_w}$ = water flux [LT^{-1}],

ρ_b = soil bulk density [ML^{-3}],

S_s = a solute reaction sink term [$ML^{-3}T^{-1}$], and

$\underset{\approx}{D}$ = $\underset{\approx}{D}(r, \theta, J_W)$ is the hydrodynamic dispersion tensor [L^2T^{-1}].

Equation 7 requires that certain assumptions be met which may not be valid under all conditions in field soils (Dagan and Bresler 1979, Gelhar and Axness 1983). However, it is the only model which is in widespread use for vadose zone chemical transport. The sink term, S_s is required to describe adsorption - desorption phenomena as well as plant uptake, chemical reaction, microbial transformations, etc. Gas phase transport, which may be important for some agricultural chemicals, is not considered in equation 7. Adsorption is often described by the first order linear kinetic rate equation:

$$S_a = \alpha (C_{a\ell} - \rho_b C_{as}) \qquad [8]$$

where S_a = the adsorption (desorption) rate [$ML^{-3}T^{-1}$],

C_{as} = the solute concentration associated with the soil matrix [M/M], and

α = a rate constant [T^{-1}].

In the root zone, transport of both water and solutes is strongly affected by temporal and spatial variability of infiltration during rainfall (or snowmelt) and by evaporation and root uptake during intervening periods. Below the root zone, gravitational forces predominate and the temporal variability of water flux decreases as depth increases.

Transport in The Groundwater Zone

The transport equations 6 and 7 simplify somewhat for groundwater flow, because the pore space remains completely filled with water. For water transport, the driving force becomes H = p + z, where p is hydrostatic pressure head, which obeys the equation

$$\underset{\approx}{\nabla} \cdot \underset{\approx}{K_s} \cdot \underset{\approx}{\nabla} H = \phi \partial p/\partial t \qquad [9]$$

where $K_s(r)$ = the saturated hydraulic conductivity tensor, and

ϕ = porosity.

The corresponding expression for chemical transport is again assumed to be the convection-dispersion equation 7. The three zones of transport are coupled by requiring continuity of water and chemical fluxes at the interfaces, and usually continuity of water potential or pressure, and chemical concentration.

Solutions of the coupled equations 1-9, with appropriate initial and boundary conditions, would provide a fairly complete description of the transport of a given chemical from the simple system depicted in figure 1. However, the equations cannot, in general, be solved analytically, so numerical techniques must be used. Furthermore, only empirical expressions are available for the critical source and sink terms in equations 4, 5 and 7, and substantial errors may occur if they are used in circumstances different from those for which the empirical relations were developed.

Finally, the tensor expressions in equations 6, 7, and 9 most certainly cannot be characterized experimentally at every point in space. Thus, they must be represented with functional forms for idealized assumptions about the nature of heterogeneity. It is now being recognized that the spatial and temporal variations in input and parameter fields not accounted for in the above formulations severely limit their applications and that little real progress will be made until these factors are addressed.

MODEL APPROACHES - SPATIAL VARIABILITY

Four approaches have been used to describe flow and transport in the presence of spatially varied physical and chemical properties:

(1) Stochastic continuum modeling. This approach uses a first- order perturbation expansion of the local flow and transport equations to derive asymptotic transport equations for the fluctuations about the mean. This technique has been applied for the transport of a dissolved species in steady groundwater flow (Gelhar et al. 1979) and to describe steady flow in the unsaturated zone (Yeh et al. 1985a,b,c). The overall goal of this type of analysis is to determine the effective large-scale behavior of the heterogeneous system and the degree of variation about the mean. For example, for transport in the saturated zone, the objective would be to determine the effective saturated conductivity and the effective diffusion-dispersion coefficient based upon knowledge of the spatial variability of the saturated hydraulic conductivity. The effective parameters so obtained can then be used in deterministic flow models, provided that the process has developed to the point where asymptotic analysis can be used.

(2) Monte Carlo analysis. In a Monte Carlo analysis, model parameters are described by distribution functions and a large number of realizations of a particular locally described process (such as the one described by equation 9) are generated by simulation to obtain the empirical distribution functions for the output variable under study. This approach has been used to describe infiltration phenomena by Smith and Hebbert (1979) and Freeze (1980). Amoozegar-Fard et al. (1982) used Monte Carlo techniques to estimate the time and space distribution of salt content in a field.

(3) Scaling theory. This approach assumes the validity of geometric similitude and derives field scale transport equations as integral ensemble averages over the frequency distribution of the scaling factor lengths. Tillotson and Nielsen (1984) reviewed methods to derive scale factors for soil properties.

(4) Parameter sensitivity analysis. In this approach we simply substitute extreme values for the various parameters in a deterministic transport model to determine the range of output values that might be observed. This analysis addresses the spatial variability problem only indirectly.

Spatial Variability and Modeling

Although several different techniques to account for spatial variability have been used and in most cases only steady-state transport of water was imposed, the overwhelming conclusion of all field studies of transport processes is that spatial variability of the key parameters must be taken into account in a realistic simulation of transport through natural soils. Indeed, more accurate conclusions may often be drawn by using a simple transport model which includes spatial variability than by using a deterministic solution of a theoretically more accurate transport model. In the past most of the studies have addressed flow and transport in the vadose or groundwater zones. A few have examined infiltration and none have addressed the effects of spatial variability

on processes involving chemical partitioning between surface runoff and infiltrating water. In light of the conceptual and analytic difficulties inherent in describing transport processes in this region, the dearth of studies is understandable. However, we will continue to have an incomplete picture of agricultural chemical transport until we do a better job of modeling the intermittent surface and near surface water, sediment and chemical transport processes.

Let us consider some of the implications of spatial variability as it might affect chemical transport governed by equations 1-9. In discussing this problem we will adopt the terminology of Dagan (1986) in referring to the scale of the flow domains: laboratory (10^{-1}-10^{0} m), local (10^{1}-10^{2} m) and regional (10^{4}-10^{5} m).

Consider an 80-ac (32.4-ha) cultivated field which lies within a 1 mi^2 (259 ha) mixed crop agricultural watershed. This field is designated as "A" in figure 2. The area of both the field and the watershed are between Dagan's (1986) local and regional scale. We assume that chemical species A, B and C are applied to the field at various times throughout the year. For example, A might be a nitrogen fertilizer, B an insecticide and C a herbicide. We are concerned with the question: "What effect will spatial and temporal variability of inputs and parameters have on predictions made by agricultural chemical transport models?"

The expected path of a chemical applied over a 1 m^2 area at point D on the field might follow the solid flow line if transported by surface runoff, and the dashed flow path (in plan view) if it travels downward through the vadose zone to groundwater. With time, the chemical species will be found in the vadose zone and in groundwater below and down gradient from the field, and intermittently

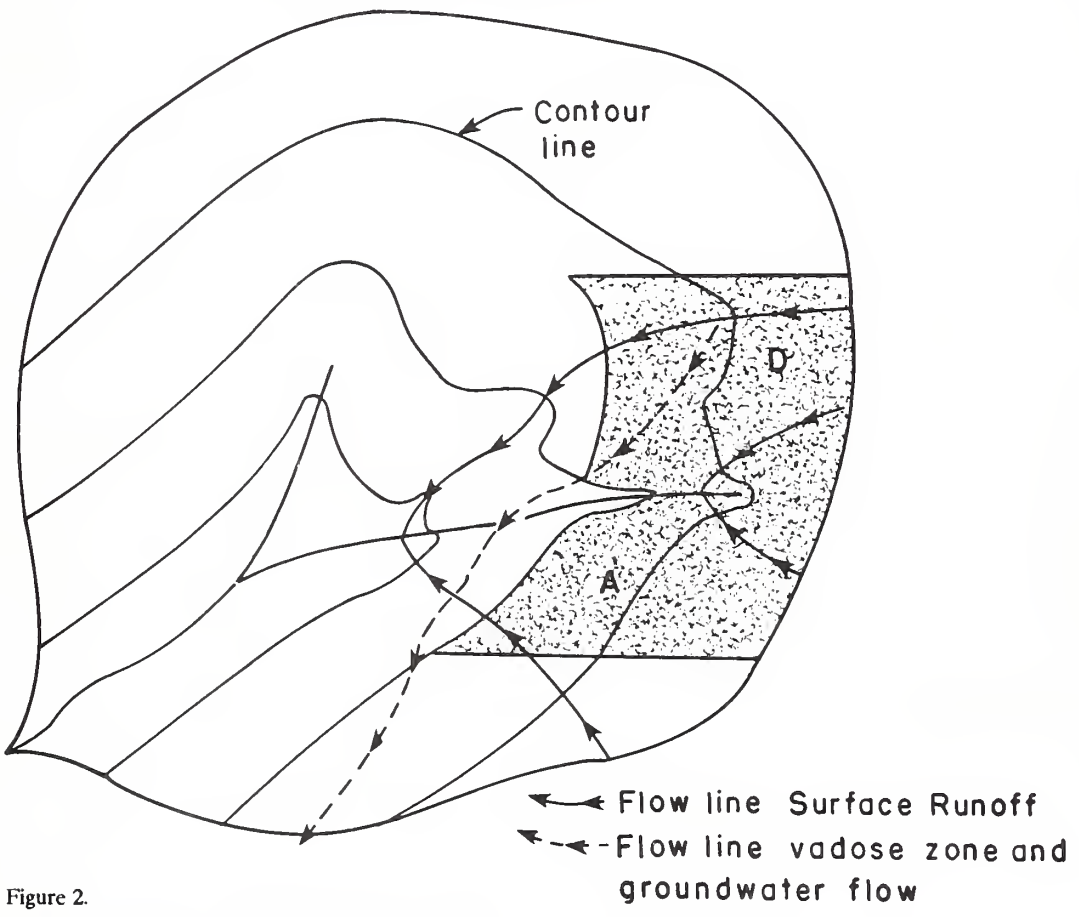


Figure 2.
Schematic drawing of field-sized chemical source area.

in surface runoff from the field which contributes to the stream. If we define a "critical" concentration, C_T, we may wish to evaluate the probability that the concentration in the ground water at specific points exceeds C_T, or the probability that this concentration is exceeded in a specific reach of stream. If the watershed system and parameters are considered to be deterministic, these probabilities depend only on the system geometry and the stochastic nature of the meteorological and chemical inputs. For real situations, however, we have incomplete information about the subsurface system geometry and the parameters can be considered as a set of random variables. The failure probabilities are therefore dependent on the joint distribution of the parameters as well as the characteristics of the input processes. As C_T approaches zero, the failure probability approaches one and we really do not need a model - the chemical in question simply should not be applied. However if there is a "safe" concentration, we need to evaluate failure probabilities and the sensitivity of these probabilities to model parameters and geometry.

Surface Transport Processes

If we consider an enlarged view of the soil surface near D during a rainstorm as shown in figure 3, we can begin to comprehend the complexities of the transport process and its dynamic nature. First we note that the surface is partitioned into two parts: the stippled areas represent portions of the surface where ponding has not occurred and the infiltration rate equals the rainfall rate. The intervening areas have ponded and the infiltration rate is spatially varying but smaller than the rainfall rate. The size of these areas depends on the current and past rainfall rates and the initial water content of the soil as well as the soil hydraulic characteristics. The spatial variability of soil hydraulic properties has been extensively documented (see Jury 1985, for a comprehensive review). The surface hydraulic properties are dynamic in that they are affected by vegetation and tillage. There is evidence that tillage treatments affect the spatial structure of infiltration as well as mean hydraulic conductivity, (Cressie and Horton 1987). The distribution of a chemical will depend on whether it is surface applied or mixed with the soil. Irrespective of the method of application, it will have significant spatial variability. Chemical located within areas that have not ponded will move downward with infiltrating water. Chemical in the intervening areas will move downward but may also be transported by surface runoff through the mechanisms described by Bailey et al. (1974). Surface microtopography and soil properties will usually be anisotropic due to soil forming processes and tillage operations.

Attempts to include spatial variability of infiltration rates date back over 20 years. The Stanford Watershed Model (Crawford and Linsley 1966) included a linear approximation to the cumulative distribution of infiltration rates within a subcatchment. The Monte Carlo approach provides useful insight into the effects of spatial variability, but is not useful in practical models. Luxmoore and Sharma (1980) subdivided a catchment into seven independent subareas. The infiltration

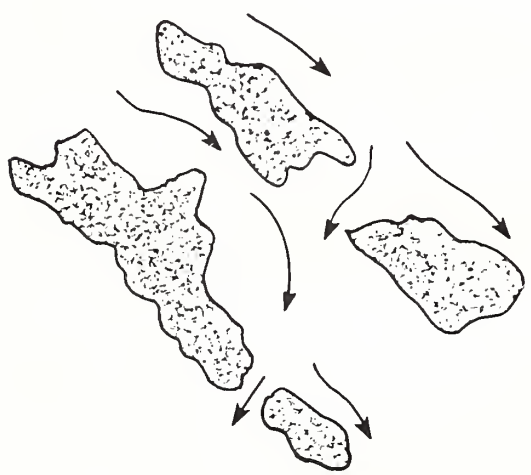


Figure 3.
Soil surface near point D during surface runoff.

characteristics of each subarea were represented by scaling factors, and the relative area was determined from the distribution of the scaling factor and was based on experimental infiltration measurements (Sharma et al. 1980). Sivapalan and Wood (1986) included the spatial correlation structure in a study of the effects of spatial heterogeneity in both soil and rainfall characteristics on the infiltration response of catchments. They concluded that the rainfall correlation structure is more important than the correlation structure of soil properties in rainfall-runoff modeling. Woolhiser and Goodrich (1988) subdivided each plane in a kinematic cascade into equal subareas with the saturated hydraulic conductivity, K_s, equal to the median value of equal probability intervals of the cumulative distribution of K_s which was assumed to be lognormal. The analogy is that each plane element in a catchment consists of independent parallel planes each with a different K_s. Outflow from each plane is added to provide either lateral inflow to a channel or the upper boundary condition for a lower plane. The cumulative distribution function (CDF) of infiltration rate for this analogy will be a step function that will vary with time and rainfall rate as shown in figure 4. The mean and coefficient of variation of K_s can vary from plane to plane to account for larger scale, position-dependent soil properties. Variations on a similar scale can be accounted for by chemical transport models such as ANSWERS (Beasley 1977), but they cannot account for small scale spatial variability. It appears that both scales are important, especially for small storms.

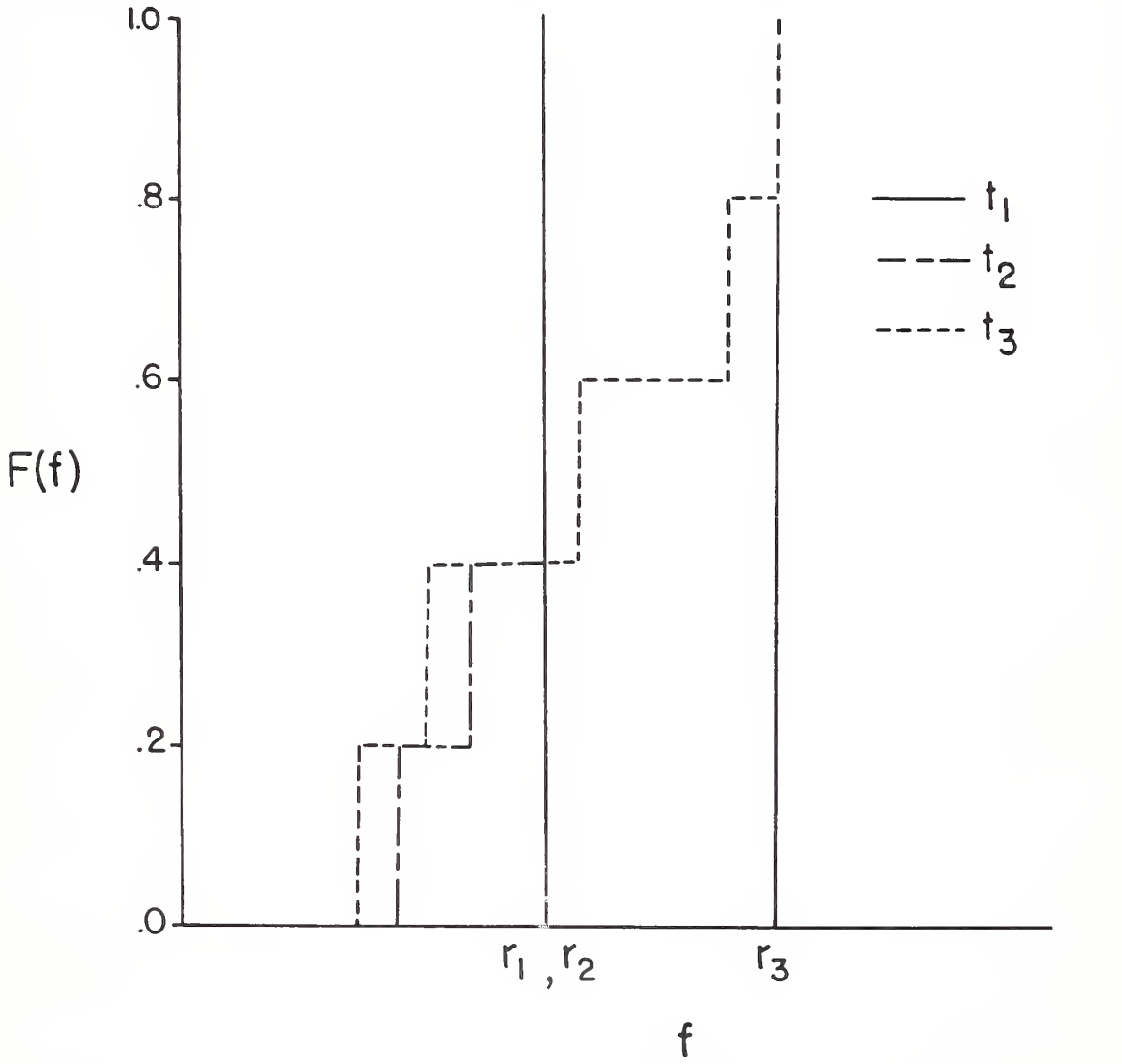


Figure 4.
Time-varying cumulative distribution function (CDF) of infiltration rates for parallel plane analogy (five equal area planes).

Existing chemical transport models essentially ignore small scale spatial variability of infiltration rates and the source term, F, in equation 4. Indeed, equations 1, 2, 4, and 5 are generally approximated by some type of mixed reactor model. For example, Huff and Kruger (1967, 1970) assumed that there was some "effective soil mass" which interacts with surface runoff. Presumably this effective soil mass consists of soil particles and aggregates being transported by overland flow as well as stationary soil particles at the soil-surface runoff interface. The combination of the effective soil mass and the surface runoff volume on a slope is treated as a completely mixed reactor. This model can be described by the equations of continuity for water, sediment and a chemical species.

The coupled ordinary differential equations are:

Continuity for water:

$$\frac{dV}{dt} = R - F - Q \qquad [10]$$

where V = the volume of water stored on the surface,
R = the integrated rainfall rate,
F = the integrated infiltration rate and
Q = the runoff rate.

Continuity for sediment:

$$\frac{dVC_s}{dt} = S_R + S_\tau + S_E + S_D - QC_s \qquad [11]$$

where the quantities on the right hand side are the same as those defined in equation 5, except that they are the integrated rates over the total slope length.

Continuity for chemical species A:

$$\frac{d}{dt}(C_A V + M_{Ae}) = RC_{AR} - FC_A - QC_A + S_A \qquad [12]$$

where M_{Ae} = the mass of species A adsorbed to the soil and sediment,
S_A = a source or sink term,
C_{AR} = the concentration of species A in the rainfall, and
C_A = the concentration in the surface runoff and the soil water of the effective soil mass.

Huff and Kruger (1970) determined the apportionment of the chemical (^{90}Sr in their study) between water and soil by a distribution coefficient, K_d:

$$K_d = \frac{f_s V}{M_s f_1} \qquad [13]$$

where f_s = fraction of the ionic species in the exchanger,
f_1 = fraction of the ionic species in the solution,
M_s = mass of the exchanger and
V = volume of solution.

Thus $M_{Ae} = K_d M_e C_A$ [14]

where M_e = effective soil mass.

Substituting equation 14 into equation 12.

$$\frac{dC_A}{dt} = \frac{RC_{AR} - C_A (F + Q) + S_A}{V + K_d M_e} \quad [15]$$

Similar approaches were used by Crawford and Donigian (1973), Donigian and Crawford (1976), Frere et al. (1980), and Novotny et al. (1978) although different expressions for the kinetics of adsorption were used in some cases.

Ingram (1979) and Ingram and Woolhiser (1980) proposed an incomplete mixing model which overcomes some of the conceptual problems associated with the complete mixing model, yet retains a relatively simple structure. Ahuja and co-workers have carried out a number of laboratory studies on the depth of rainfall-runoff-soil interaction and the mechanisms of transfer of surface-applied chemicals to runoff, Ahuja (1982), Ahuja et al. (1981), Sharpley et al. (1981), Ahuja et al. (1982), Ahuja and Lehman (1983), Heathman et al. (1985), and Sharpley (1985). They found that the effective depth of interaction is affected by rainfall intensity, slope, soil aggregation, mulch and other factors and that the completely mixed reactor model is often not a good approximation. Ahuja and Lehman (1983) suggest that upward chemical transport near the soil surface might be represented as an accelerated diffusion process.

Emmerich et al. (1989) have examined the distortions induced by lumping the chemical transport model by comparing solutions from a lumped model to those from an advective transport model. They showed that the lumped formulation generally leads to damping of peak concentrations and poor estimates of pollutant arrival times if the chemical is applied to only part of the surface.

Vadose Zone Transport

Spatial variations in soil water transport and retention parameters have been addressed in several transport and flow models. Warrick et al. (1977) used Monte Carlo simulations to represent the effects of spatial variability on drainage rate. They found that the field-averaged drainage rate could not be represented with a one dimensional model which used average values for the transport parameters. A similar result was found by Amoozegar-Fard et al. (1982), who simulated chemical transport by the Monte Carlo method and found that the field scale dispersion was dominated by convective velocity differences and not by local dispersion. Kool et al. (1987) provide a review of parameter estimation techniques for unsaturated flow and transport models.

A different approach was taken by Jury (1982), who proposed a transfer function model of solute travel times to simulate solute leaching through the vadose zone. A stochastic form of this model which assumes zero lateral mixing was shown to describe solute movement adequately over the top two meters of soil (Jury et al. 1982). A later test of this model described solute movement over 0.64 ha accurately to a depth of 25 m after a calibration at 30 cm, while the convection-dispersion model (eq. 9) was unsuccessful in the same test (Butters and Jury 1989).

Dagan and Bresler (1983) point out that although the determination of the distribution of moisture content in three spatial dimensions and time in a heterogeneous field is a "computational nightmare", estimates of the average as a function of depth and time is relatively simple. They assumed that horizontal flow components are negligible and that K_s was lognormally distributed. They divided the CDF of log K_s into equal probability classes and simulated steady, vertical water flow with a piston type model. They found that this simplified procedure led to good approximations of the means and variances of water content, provided that the field is sufficiently heterogeneous (Bresler and Dagan 1983a). The integral scale of log K_s must also be much smaller

than the length scale characterizing the field. For unsaturated flow they (and others) found that the concept of equivalent uniform soil is not valid. In the third paper of the series, Bresler and Dagan (1983b), used a closed form solution of the convective dispersion equation with convective velocities determined by their simplified water transport model to calculate means and variances of salt concentration over a field as functions of time. Again they found that the approximations led to good results where spatial variability was large, and concluded that further refinement of models of transport in homogeneous columns does not seem warranted but that research should be concentrated on field variability.

A scaling theory model of solute transport was proposed by Dagan and Bresler (1979) in which the local scaling length is treated as a random variable. This model has not received any experimental tests.

Transport in Groundwater

The most promising approaches used to model the spatial variability of solute and water movement in groundwater have used stochastic continuum formulations of the flow process. In this method the solute and water parameters which manifest substantial spatial variability are treated as random functions and the local transport equation is expanded to first order in the fluctuations of the parameters about their mean values. Gelhar and Axness (1983) used this approach to derive an asymptotic form of the solute macrodispersion coefficient which was a function of the fluctuations in local hydraulic conductivity and the degree of lateral mixing of solute concentration. Recently, Dagan (1984, 1987) has extended this approach to create a continuous description of solute movement. This model reproduced the key features of a solute pulse migrating over a long distance in a groundwater flow experiment (Freyberg 1986, Roberts et al. 1986, Mackay et al. 1986, Sudicky 1986).

Gelhar (1986) discussed research developments treating groundwater flow in a probabilistic framework. He pointed out that although considerable progress has been made in the stochastic analysis of subsurface flow, applications have been limited. Major reasons are that many cases of field transport are fundamentally three dimensional and in many cases boundary conditions are poorly defined because few data are available. Gelhar (1986) pointed out that spatial correlation scales should not be viewed in any absolute sense, but will depend on the scale of the problem at hand. However for the perturbation approach to work, the correlation scale must be small relative to the scale of the problem. The perturbation approach has been quite robust for steady-state saturated problems, but it remains to be seen if it is adequate for more complicated phenomena, including unsaturated flow and solute transport.

Recently, some investigators have studied the spatial variability of solute concentrations in the root zone and in the vadose zone near the water table. Ronen et al. (1987) used a multilayer dialysis cell sampling device and found large variations in the concentration of Cl^-, NO_3^- and SO_4^{2-} in the upper water layers of a polluted aquifer. This variation was at the microscale level (3 cm) and raises questions about current sampling procedures.

Basin Scale Transport

On the basin scale, Rinaldo and Marani (1987) have suggested that because of the complexity of transport processes in heterogeneous media and the scarcity of data, that simple models aimed at describing dominant modes of behavior are preferable to distributed models of microprocesses. "Not only, in fact, might large-scale simulation of all microprocesses prove impossible, but it may also be unnecessary." They developed general expressions for basin mass response functions of solute transport in a probabilistic framework drawing on the earlier work on runoff (Lienhard 1964, Rodriguez-Iturbe et al. 1979), sediment yield (Moore 1984), and solute transport in porous media (Jury 1982, Jury et al. 1986, Sposito et al. 1986).

MODEL APPROACHES - TEMPORAL VARIABILITY

Due to the stochastic nature of the driving meteorologic inputs, delivery of dissolved and adsorbed chemicals to receiving waters by surface runoff and erosion exhibits extreme temporal variability. Although flow processes in the vadose and groundwater zones tend to damp out these temporal fluctuations, substantial annual variability in chemical transport frequently exists.

The most common method of dealing with temporal variability is through the simulation approach. By using long sequences of measured or simulated meteorologic input data to drive a chemical transport model it is possible to estimate the CDF of annual chemical transport in surface runoff from a field or out of a root zone. In general, we must allow a finite probability of zero transport, although this probability will be small in humid regions. Small and Mular (1987) developed an analytical model to assess the effects of meteorological temporal variability on the fraction of contaminant that escapes a zone of reaction. They assumed that the infiltration process was Poisson with a gamma distribution of amounts and used a one dimensional advective diffusion equation for transport with a first order decay model to predict the distribution function of the breakthrough fraction. Spatial variability was ignored. They compared their results with results from a Monte Carlo model, and found that their model underestimated dispersion.

DISCUSSION

At the present stage of agricultural water quality model development, the typical model consists of a heterogeneous assembly of elements, some very detailed and possibly including stochastic components to deal with spatial variability, other elements highly simplified and treating the process only with respect to its overall behavior (and other components may be missing entirely). Predictions based upon these models are subject to substantial errors due to uncertainties about model structure, parameter values and inputs. For an excellent review on this topic see Beck (1987). It is apparent that spatial and temporal variability exist over a wide range of scales. Dagan (1986) points out that the computational scale and the measurement scales are also very important. In developing a mathematical model to describe transport phenomena at the laboratory scale, the computational length and time scales can be the same or smaller than the appropriate length and time scales for the prototype system and models can give quite good results. However at the field scale, the computational length scale can be much larger than the heterogeneity scales for chemical application, infiltration parameters, etc. If large scale spatial averages are used for inputs and parameters, the results are often severely biased. Therefore, explicit procedures must be devised to incorporate small scale spatial variability without excessive computational effort. The work of Bresler and Dagan (1983b) for salt transport in the vadose zone is an example. This approach should be attempted for solute transport in surface runoff, possibly using the parallel plane analogy of Woolhiser and Goodrich (1988). Such a model would allow differential chemical transport from a field, with varying portions of the field contributing runoff depending on rainfall intensity. The large scale variations in the mean saturated hydraulic conductivity can already be taken into account by distributed models.

There has been little work in evaluating the spatial characteristics of chemical inputs, and small scale variability is apparently not accounted for in existing models, although it certainly exists.

As we move from field to watershed scale, the computational scale will necessarily increase as well. However there are upper limits to the computational time scale. Woolhiser (1986) suggested that a five minute sampling time is the maximum that should be used to accurately estimate runoff generated by the Hortonian mechanism. As the computational length scale increases, it may be necessary to increase the coefficient of variation of log K to accurately model small storms. The maximum size of the spatial computational scale is probably governed by areal characteristics of precipitation.

LITERATURE CITED

Ahuja, L.R. 1982. Release of soluble chemical from soil to runoff. Transactions of the American Society of Agricultural Engineers 25:948-953, 960.

Ahuja, L.R., and O.R. Lehman. 1983. The extent and nature of rainfall-soil interaction in the release of soluble chemicals to runoff. Journal of Environmental Quality 12:34-40.

Ahuja, L.R., A.N. Sharpley, and O.R. Lehman. 1982. Effect of soil slope and rainfall characteristics on phosphorus in runoff. Journal of Environmental Quality 11:9-13.

Ahuja, L.R., A.N. Sharpley, M. Yamamoto, and R.G. Menzel. 1981. The depth of rainfall-runoff-soil interaction as determined by ^{32}P. Water Resources Research 17:969-974.

Akan, A.O., and B.C. Yen. 1981. Mathematical model of shallow water flow over porous media. Journal of the Hydraulics Division, American Society of Civil Engineers, Proceeding Paper 16213, 107:479-494.

Amoozegar-Fard, A., D.R. Nielsen, and A.W. Warrick. 1982. Soil solute concentration distributions for spatially varying pore water velocities and apparent diffusion coefficients. Soil Science Society of America Journal 46:3-9.

Bailey, G.W., R.R. Swank, and H.P. Nicholson. 1974. Predicting pesticide runoff from agricultural land: A conceptual model. Journal of Environmental Quality 3:95-102.

Beasley, D.B. 1977. ANSWERS: A mathematical model for simulating the effects of land use and management on water quality. Ph.D. Thesis. Purdue University, West Lafayette, IN.

Beck, M.B. 1987. Water quality modeling: A review of the analysis of uncertainty. Water Resources Research 23:1393-1442.

Bresler, E., and G. Dagan. 1983a. Unsaturated flow in spatially variable fields 2. Application of water flow model to various fields. Water Resources Research 19:421-428.

Bresler, E., and G. Dagan. 1983b. Unsaturated flow in spatially variable fields 3. Solute transport models and their application to two fields. Water Resources Research 19:429-435.

Butters, G.L., and W.A. Jury. 1989. Field scale transport of bromide through unsaturated soil. Water Resources Research (In press).

Crawford, N.H., and A.S. Donigian. 1973. Pesticide transport and runoff model for agricultural lands. EPA-660/2-74-013. U.S. Environmental Protection Agency, Washington, D.C.

Crawford, N.H., and R.K. Linsley. 1966. Digital simulation in hydrology: Stanford Watershed Model IV. Department of Civil Engineering, Stanford University, Stanford, CA. Technical Report No. 39. 210 pp.

Cressie, N.A.C., and R. Horton. 1987. A robust-resistant spatial analysis of soil water infiltration. Water Resources Research 23:911-917.

Dagan, G. 1984. Solute transport in heterogeneous porous formations. Journal of Fluid Mechanics 145:151-177.

Dagan, G. 1986. Statistical theory of groundwater flow and transport: Pore to laboratory, laboratory to formation, and formation to regional scale. Water Resources Research 22:120S-134S.

Dagan, G. 1987. Theory of solute transport by groundwater. Annual Review of Fluid Mechanics 19:183-215.

Dagan, G., and E. Bresler. 1979. Solute dispersion in unsaturated heterogeneous soil at field scale. Soil Science Society of America Journal 43:461-467.

Dagan, G., and E. Bresler. 1983. Unsaturated flow in spatially variable fields 1. Derivation of models of infiltration and redistribution. Water Resources Research 19:413-420.

Donigian, A.S., and N.H. Crawford. 1976. Modelling pesticides and nutrients on agricultural lands. EPA-600/2-76-043, U.S. Environmental Protection Agency, Athens, GA.

Emmerich, W.E., D.A. Woolhiser, and E.D. Shirley. 1989. Comparison of lumped and distributed models for chemical transport by surface runoff. Journal of Environmental Quality 18:120-126.

Freeze, R.A. 1980. A Stochastic-conceptual analysis of rainfall-runoff processes on a hillslope. Water Resources Research 16:391-408.

Frere, M.H., J.D. Ross, and L.J. Lane. 1980. The nutrient submodel. Chapter 4 In W.G. Knisel, ed. CREAMS: A field scale model for chemicals, runoff and erosion from agricultural management systems. USDA Conservation Research Report 26:65-87.

Freyberg, D.L. 1986. A natural gradient experiment on solute transport in a sand aquifer, 2. Spatial moments and the advection and dispersion of nonreactive tracers. Water Resources Research 22:2031-2046.

Gelhar, L.W. 1986. Stochastic subsurface hydrology from theory to application. Water Resources Research 22:135S-145S.

Gelhar, L.W., and C. Axness. 1983. Three dimensional stochastic analysis of macrodispersion in aquifers. Water Resources Research 19:161-180.

Gelhar, L.W., A.L. Gutjahr, and R.N. Naff. 1979. Stochastic analysis of macrodispersion in a stratified aquifer. Water Resources Research 15:1387-1397.

Heathman, G.C., L.R. Ahuja, and O.R. Lehman. 1985. The transfer of soil surface-applied chemicals to runoff. Transactions of the American Society of Civil Engineering 28:1909-1915, 1920.

Huff, D.D., and P. Kruger. 1967. A numerical model for the hydrologic transport of radioactive aerosol from precipitation to water supplies. Isotope Techniques. In The Hydrologic Cycle, American Geophysical Union, Geophysics Monograph Series No. 11:85-96.

Huff, D.D., and P. Kruger. 1970. Simulation of the hydrologic transport of radioactive aerosols. Advances in Chemistry No. 93. Radionuclides in the Environment: 487-505.

Ingram, J.J. 1979. Chemical transfer from a saturated soil into overland flow. MS Thesis, Colorado State University, Fort Collins, CO. 148 pp.

Ingram, J.J., and D.A. Woolhiser. 1980. Chemical transfer into overland flow. In Proceedings A.S.C.E. Symposium on Watershed Management. Boise, ID. 40-53.

Jury, W.A. 1982. Simulation of solute transport using a transfer function model. Water Resources Research 18:363-368.

Jury, W.A. 1985. Spatial variability of soil physical parameters in solute migration: A critical literature review. Electric Power Research Institute Interim Report EA-4228, 80 pp. Palo Alto, CA.

Jury, W.B., G. Sposito, and R.E. White. 1986. A transfer function model of solute transport through soil. 1. Fundamental concepts. Water Resources Research 22:243-247.

Jury, W.A., L.H. Stolzy, and P. Shouse. 1982. A field test of the transfer function model for predicting solute transport. Water Resources Research18:369-375.

Kool, J.B., J.C. Parker, and M. TH. Van Genuchten. 1987. Parameter estimation for unsaturated flow and transport models - a review. Journal of Hydrology 91:255-293.

Lienhard, J.H. 1964. A statistical mechanical prediction of the dimensionless unit hydrograph. J. Geophys. Res. 69:5231-5238.

Luxmoore, R.J., and M.L. Sharma. 1980. Runoff response to soil heterogeneity: Experimental and simulation comparisons for two contrasting watersheds. Water Resources Research 16:675-684.

Mackay, D.M., D.L. Freyberg, P.V. Roberts, and J.A. Cherry. 1986. A natural gradient experiment on solute transport in a sand aquifer. 1. Approach and overview of plume movement. Water Resources Research 22:2017- 2029.

Moore, R.J. 1984. A dynamic model of basin sediment yield. Water Resources Research 20:89-103.

Novotny, V., H. Tran, G.V. Simsimon, and G. Chesters. 1978. Mathematical modeling of land runoff contaminated by phosphorus. J. Water Pollu. Control Fed. 50:101-112.

Rinaldo, A., and A. Marani. 1987. Basin scale model of solute transport. Water Resources Research 23:2107-2118.

Roberts, P.V., M.N. Goltz, and D.M. Mackay. 1986. A natural gradient experiment on solute transport in a sand aquifer 3. Retardation estimates and mass balances for organic solutes. Water Resources Research 22:2047- 2058.

Rodriguez-Iturbe, I., G. Devoto, and J.B. Valdes. 1979. Discharge response analysis and hydrologic similarity: the interrelation between the geomorphic IUH and storm characteristics. Water Resources Research 15:1435-1444.

Ronen, D., M. Magaritz, H. Gvirtzman, and W. Garner. 1987. Microscale chemical heterogeneity in groundwater. Journal of Hydrology 92:173-178.

Sharma, M.L., G.A. Gander, and C.G. Hunt. 1980. Spatial variability of infiltration in a watershed. Journal of Hydrology 45:101-122.

Sharpley, A.N. 1985. Depth of surface soil-runoff interaction as affected by rainfall, soil slope, and management. Soil Science Society of America Journal 49:1010-1015.

Sharpley, A.N., L.R. Ahuja, and R.G. Menzel. 1981. The release of soil phosphorus to runoff in relation to the kinetics of desorption. Journal of Environmental Quality 10:386-391.

Sivapalan, M., and E.F. Wood. 1986. Spatial heterogeneity and scale in the infiltration response of catchments. In V.K. Gupta, I. Rodriguez-Iturbe, and E.F. Wood eds. Scale Problems in Hydrology. 81-106. D. Reidel, Dordrecht, The Netherlands.

Small, M.J., and J.R. Mular. 1987. Long term pollutant degradation in the unsaturated zone with stochastic rainfall infiltration. Water Resources Research 23:2246-2256.

Smith, R.E., and R.H.B. Hebbert. 1979. A Monte Carlo Analysis of the Hydrologic Effects of Spatial Variability of Infiltration. Water Resources Research 15:419-429.

Sposito, G., R.E. White, P.R. Darrah, and W.A. Jury. 1986. A transfer function model of solute transport through soil. 3. The convection-dispersion equation. Water Resources Research 22:255-262.

Sudicky, E.A. 1986. A natural gradient experiment on solute transport in a sand aquifer: Spatial variability of hydraulic conductivity and its role in the dispersion process. Water Resources Research 22:2069-2082.

Tillotson, P.M., and D.R. Nielsen. 1984. Scale factors in soil science. Soil Science Society of America Journal 48:953-959.

Warrick, A.W., G.J. Mullen, and D.R. Nielsen. 1977. Predictions of soil water flux based upon field measured soil water properties. Soil Science Society of America Journal 41:14-18.

Woolhiser, D.A. 1986. Sensitivity of calculated peak runoff rates to rainfall-sampling frequency. Integrated Design of Hydrologic Networks (Proceedings of the Budapest Symposium). International Association of Hydrologic Sciences, Publication 158:161-171.

Woolhiser, D.A., and D.C. Goodrich. 1988. Effect of Storm Rainfall Intensity Patterns on Surface Runoff. Journal of Hydrology 102:335-354.

Yeh, T.-C.J., L.W. Gelhar, and A.L. Gutjahr. 1985a. Stochastic Analysis of Unsaturated Flow in Heterogeneous Soils. 1. Statistically Isotropic Media. Water Resources Research 21:447-456.

Yeh, T.-C.J., L.W. Gelhar, and A.L. Gutjahr. 1985b. Stochastic Analysis of Unsaturated Flow in Heterogeneous Soils. 2. Statistically Anisotropic Media with Variable *alpha. Water Resources Research 21:457-464.

Yeh, T.-C.J., L.W. Gelhar, and A.L. Gutjahr. 1985c. Stochastic Analysis of Unsaturated Flow in Heterogeneous Soils. 3. Observations and Applications. Water Resources Research 21:465-471.

EFFECT OF SPATIAL VARIABILITY UPON SUBSURFACE TRANSPORT OF SOLUTES FROM NONPOINT SOURCES

G. Dagan[1], D. Russo[2] and E. Bresler[2]

ABSTRACT

The review discusses the effect of spatial variability of soil hydraulic properties on transport of solutes in the subsurface, through the mechanism of convection by water flow. The general ideas are: (i) heterogeneity is of an irregular spatial variation and is regarded as random, (ii) the correlation scale of the relevant properties (mainly, saturated hydraulic conductivity) is much larger than the pore scale, (iii) the spatial variation of the flow velocity caused by heterogeneity results in an irregular pattern of solutes distribution, which also is regarded as random, and in the enhancement of their spreading, and (iv) in the case of agricultural nonpoint sources the input area of the solutes is generally large compared to the conductivity correlation scale; consequently, the space average of the concentration over such an area is subject to a much lesser degree of uncertainty than the point value. The major objective of modeling transport is seen as deriving partial differential equations for the space-averaged concentration. In these equations the effect of heterogeneity manifests in an increased and time-dependent effective dispersion coefficient. The advances in modeling transport through the unsaturated zone and by groundwater are reviewed separately, along the above general lines.

INTRODUCTION

Agricultural activities constitute one of the wide-spread sources of pollution of the subsurface environment, i.e. of the upper soil layer and of groundwater, and the prediction of the fate of pollutants is a subject of paramount importance. In this review we concentrate on computational methods of the transport process. Such methods are based on a conceptual framework, on subsequent theoretical developments and ultimately on particular models adapted to the problem at hand. The purpose of these developments is two-fold: to advance the basic understanding of the transport processes, and to provide quantitative predictive tools. In turn, these tools are to be used to predict future distribution of pollutants in the subsurface and for management purposes, i.e. to ameliorate the soil and water quality. Our interest resides in transport occurring at the field scale, on a much larger scale than the laboratory's. The review focuses on some recent developments of the theory and of models of transport at formation scale. One of the distinctive features of the soil at field scale is its heterogeneity, or spatial variability of its properties which affect transport. This variability is generally of an irregular fashion and occurs on a scale which is not captured by laboratory samples. These features have a definite effect on the spatial distribution of solutes, as a result of transport through the heterogeneous medium. The detailed measurement of concentrations in the field is expensive and time-consuming, and there are not many examples of elaborate field surveys in the literature. To illustrate the complexity and irregular distribution of solutes caused by field-scale heterogeneity, we reproduce in figures 1, 2 and 3 the outcome of three recent experiments, the first two pertaining to solute bodies transported by groundwater and the third to transient infiltration in the unsaturated zone of the upper soil layer. In view of these findings, it is clear that any meaningful prediction of the fate of pollutants in the field has to account for spatial variability of properties and processes which affect transport, and this is precisely the aim of the

[1]G. Dagan, Faculty of Engineering, Tel-Aviv University, Tel Aviv, Israel

[2]D. Russo and E. Bresler, Institute of Soil and Water, Agricultural Research Organization, Bet Dagan, Israel.

paper. In the second section we shall dwell upon some fundamental aspects of the emerging theory of transport in heterogeneous formations. The final sections describe a few recent developments of modeling the unsaturated zone and groundwater, respectively.

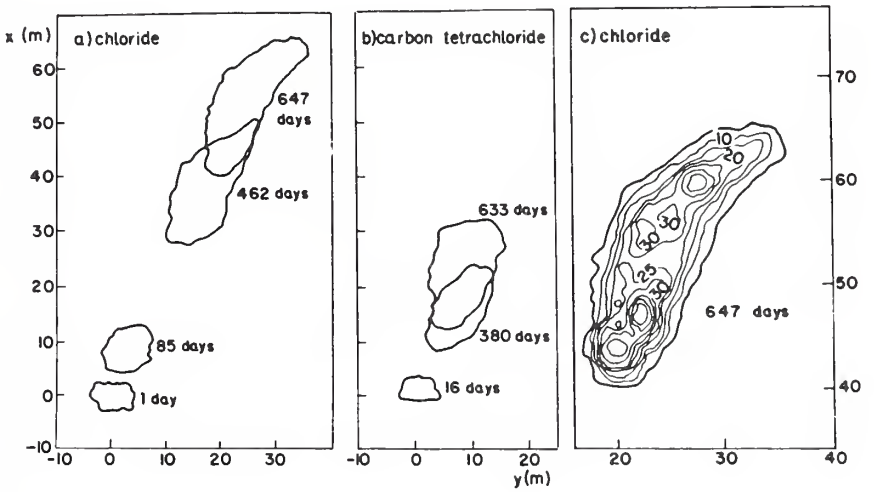


Figure 1.
Contours of solute body in the horizontal plane in the Borden-site field experiment; (a) an inert (b) a reactive tracer and (c) a blow-up of the distribution of the vertically averaged chloride concentration (reproduced from Mackay et al. 1986b).

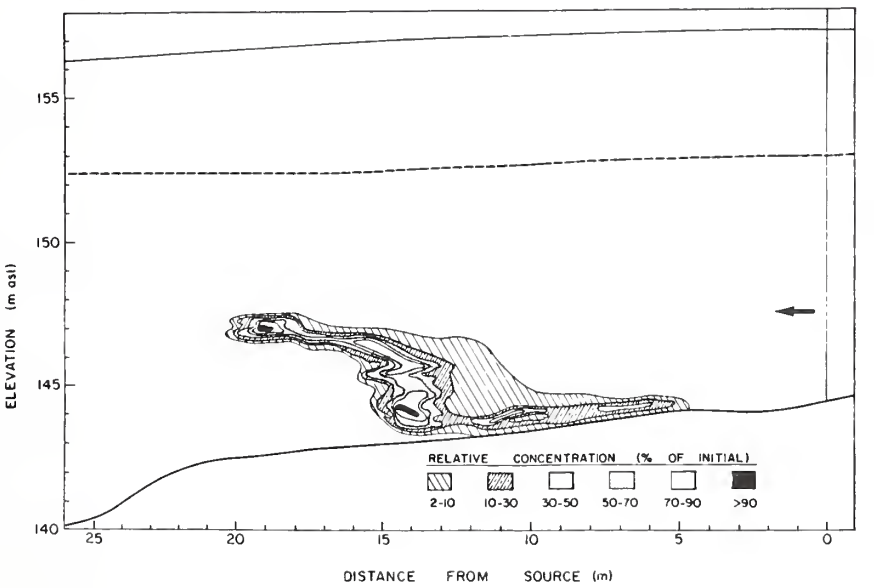


Figure 2.
Schematic cross-sectional view of the concentration of a non-reactive tracer in the field test at the Chalk River Nuclear Laboratories (reproduced from Moltyaner 1987).

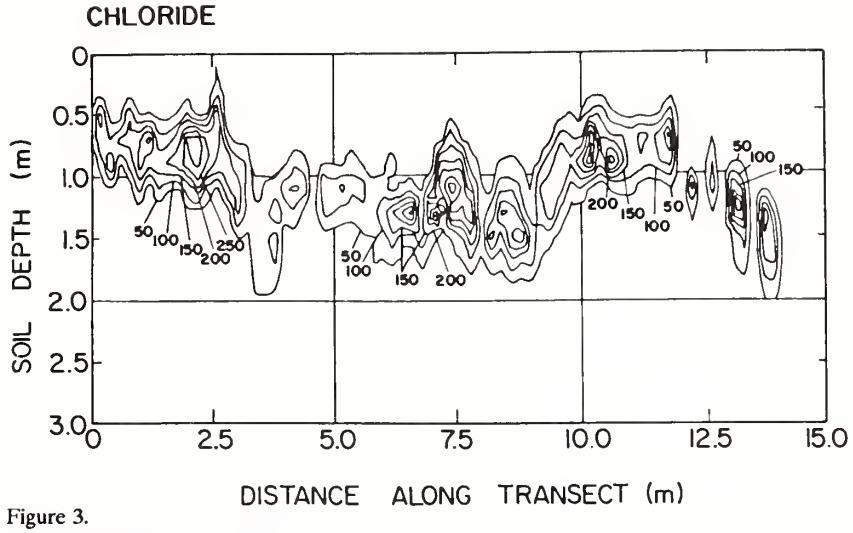


Figure 3.
Observed distributions of the gravimetric chloride concentration in unsaturated soil (reproduced from Schulin et al. 1987).

FUNDAMENTAL ASPECTS

The Traditional Approach (Laboratory Scale)

The material making up the subsurface, the soil, is a highly complex medium due to its complicated micro-geometry at the pore scale, and the intricate distribution of its physical and chemical properties. Since in applications one is interested in the behavior of samples of soil which are large compared to the pore scale, the main aim of the theory of flow and transport was to develop macroscopic relationships between the variables and parameters characterizing the system. This is the content of the traditional literature of soil physics, of theory of saturated flow and of reservoir engineering, and is still the object of ongoing research. While the starting point is from the basic laws of physics, the validation and experimental evidence underlying the macroscopic equations are traditionally achieved by controlled laboratory experiments, the most common device being the ubiquitous soil column. Such a column is characterized by a length scale of the order of tens of centimeters, which is much larger than the pore scale, and it is generally homogeneous, in the sense that the gross features of the pore structure repeat themselves throughout the sample. This scale has been termed "the laboratory scale" in the review by Dagan (1986), along which the present discussion is oriented. The basic macroscopic laws governing flow and transport, expressed by differential equations, have been summarized in text books (e.g. Bresler et al. 1982). The following is a succinct list.

The macroscopic equation of conservation of water volume, regarded as an incompressible fluid, is

$$\frac{\partial \theta}{\partial t} + \nabla \mathbf{q} = -r \qquad [1]$$

where θ = the water content (volume of water per volume of medium),
∇ = the divergence differential operator,
$\mathbf{q}$ = the specific discharge (volume per unit area of medium and unit time) and
r = a sink term related, for instance, to the presence of roots.

The flux-force relationship is expressed by Darcy's law

$$\mathbf{q} = -K \nabla(-\psi+z) \quad [2]$$

where K = the hydraulic conductivity,
ψ = the matric potential and
z = the elevation.

For saturated flow in an inert matrix, as it is generally the case for aquifer flow, $K=K_s$ is a macroscopic property of the medium. The matters are more complicated for unsaturated soils in which K depends on θ and on the solute concentration C (mass of solute per volume of water), especially in the presence of clay minerals and of organic fractions. The matric potential can be related to the water pressure p by $\psi=-p/\gamma$, where γ is the specific weight and ψ is a hysterical function of θ in unsaturated state. In saturated flow the water head $\phi=-\psi+z$ serves as a potential in 2. In the case in which K is independent of C, equations 1 and 2 form a closed system, once $K(\theta)$ and $\psi(\theta)$, macroscopic relations characterizing the particular sample, are given, and elimination of $\mathbf{q}$ yields Richards' equation.

The equation of conservation of solute mass can be written in a general form as follows

$$\frac{\partial(\theta C)}{\partial t} + \nabla(\mathbf{q}C) = \nabla(\theta \underline{\mathbf{D}} C) - S \quad [3]$$

where $\underline{\mathbf{D}}$ = the pore-scale dispersion tensor and
S = a sink term related to chemical decay and adsorption by the matrix.

In the latter case one may write

$$S = \frac{\partial}{\partial t}\left[(1-\theta)\frac{\partial C_s}{\partial t}\right] \quad [4]$$

where C_s is the mass of solute per unit volume of matrix (in some interpretations it also includes the so called "immobile water"). An additional relationship, which must be given, relates C and C_s. Among the many substances which may move with water in soil, we limit the discussion here to solutes at sufficiently low concentrations, such that they do not significantly affect the density or viscosity of the solution.

The "phenomenological" relationships $K(\theta,C)$, $\psi(\theta,C)$, $S(C,C_s)$ as well as the dependence of $\underline{\mathbf{D}}$ upon $\mathbf{q}$ and θ have been the object of extensive investigations, theoretical and experimental, whose review is beyond our scope. Eventually, the dependence of K and ψ on θ is reduced to simple relationships which include a few parameters, like K_s and θ_s, the conductivity and water content at saturation, θ_{ir}, the irreducible θ, as well as a few coefficients related to the pore-size distribution. Similarly, $\underline{\mathbf{D}}$ can be assumed to depend linearly on $\mathbf{q}/\theta$, by two coefficients, λ_L and λ_T, the longitudinal and lateral, dispersivities. For a slow sorption process, following a linear isotherm, the term S in equations 3 and 4 manifests in a retardation coefficient appearing in front of equation 3. These relationships become much simpler in the case of aquifer saturated flow.

After establishing the differential equations of flow and transport, based on conceptualization of the various processes and on physical principles (for recent reviews, see for instance Sposito et al. 1986, Nielsen et al. 1986), the next step in the traditional approach is to solve them with appropriate initial and boundary conditions. This has been the object of a rich literature, of a mathematical nature, covering analytical and numerical methods. Many of these solutions and their applications are reviewed in papers presented at this meeting. The mathematical difficulties are significant, due to the nonlinear character of equations 1-3, and various simplifications have been suggested in the past. As an example, it is common to regard transport in the unsaturated

zone as vertical, i.e. depending solely on z and t , whereas in aquifer flow the average head gradient is viewed as horizontal.

Traditionally, solutions of the transport equations have been applied to the field, with the usual assumption being that the situation prevailing in the laboratory, i.e. spatial homogeneity of the various macroscopic properties, is valid at field scale as well. As for the values of these properties, some average figures have been selected as representative. This picture has been under scrutiny recently, and is the topic of the following discussion.

Field Spatial Variability

Although the spatial variability of macroscopic properties in the field has been recognized as omnipresent for a long time, it is only in the last decade that this subject has been investigated in a systematic manner in a theoretical framework (for recent reviews, see Neuman 1982, Dagan 1986 and Gelhar 1986). First, to illustrate the type of heterogeneity one may encounter in an aquifer, we have reproduced in figure 4 a mapping of hydraulic conductivities determined by a thorough sampling, at the Borden site, the same one which pertains to figure 1. Again, the irregular, seemingly erratic, spatial variation of K_s is evident in this picture. This variability is not captured by any single laboratory sample, which is locally homogeneous. The fundamental questions addressed by recent studies were: 1) What is the impact of the spatial variability on flow and transport at the field scale? 2) How can one incorporate these effects in predictive models? 3) How are management policies and data collection strategies affected by these findings? In the following we shall review briefly a few advancements in this area.

It is an accepted tenet that spatial changes of soil properties are subject to uncertainty, due to (i) the large range of values encountered in the same formation, (ii) to irregular variation and (iii) to the limited number of measurements available in practice. The appropriate mathematical setting to represent uncertainty is probability theory and variables like K_s, θ_s, and λ_L, are regarded as random space functions. In other words, such a variable is a function of the space coordinate x, and its value at a point is a random variable which is defined in terms of its various statistical moments, e.g. expected value and variance. When we talk about the value at a point, we refer to the property measured by a device which has a dimension large compared to the pore-scale and whose centroid is at x. In other words, from now on, our variables are the macroscopic ones appearing in equations 1-4, and the underlying porous micro-structure is of no direct concern. In this stochastic model, the actual formation is regarded as one realization of an ensemble of formations, which serves to define various statistical moments. This concept poses an immediate problem, since in reality the only available realization is the actual one. This conceptual difficulty can be removed if we relate to the ensemble as a set of "phantom" formations, with the same statistical distributions of properties as the actual one, being reminded that we do not know the detailed picture in the actual field, anyway. The argument is circular, however, unless some sort of probabilistic stationarity prevails, so that by the methods of statistical inference one can determine the moments of interest from the available realization. The assumption of stationarity is adopted in this theory, and although it may seem quite arbitrary, it can be, and it has been, validated by various a-posteriori checks, e.g. by using maximum likelihood procedures.

Hence, in the simplest representation, spatially variable properties are regarded as stationary random functions, which are characterized at second-order by the expected value (which may have a drift) and by a two-point covariance. To illustrate the point, it was found by analyzing many measurements (e.g. see Freeze 1975) that $Y = \ell n\, K_s$ is normal, and an acceptable form for its covariance is the exponential one

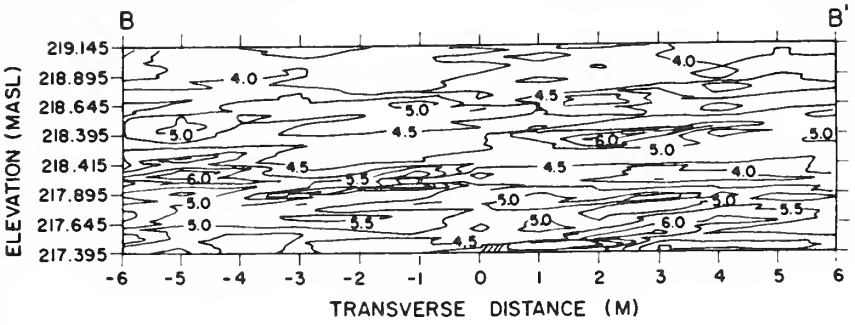


Figure 4.
Example of vertical hydraulic log-conductivity profiles determined from cores taken at 1-m horizontal distance at the Borden tracer test site (from Sudicky 1986; conductivity unit is cm/sec).

$$C_Y(x_1,x_2) = w[1 - H(|x|)] + -\sigma_Y^2 \exp[-|x_1-x_2|/I_Y] \qquad [5]$$

where
- x_1,x_2 = the coordinate vectors of two arbitrary points,
- H = the Heaviside step-function ($H=0$ for $x<0$, $H=1$ for $x>0$),
- σ_Y^2 = the variance of spatially correlated residuals,
- $w+\sigma_Y^2$ = the total variance, and
- I_Y = the correlation scale, indicative of the distance over which the values of Y are correlated.

Thus, the statistical structure of the lognormal conductivity is encapsuled by four parameters: the expected value, m_Y, the variance of uncorrelated residuals, w (the so called nugget effect), the variance, σ_Y^2, and the correlation scale, I_Y. The nugget represents either measurement errors or erratic variations at a smaller scale than the minimal distance between measurement points. The identification of these parameters from measurements with the aid of a maximum likelihood procedure is explored, for instance, by Hoeksema and Kitanidis (1984).

A complete statistical characterization at second-order of the formation, implies the determination of similar sets of parameters for other properties which affect flow and transport. This may be a formidable task in terms of required measurements, and one may have to rely on prior information from similar formations. Some results concerning K_s are summarized by Gelhar (1986), and an illustration of analysis of parameters characterizing the $K(\theta)$ and $\psi(\theta)$ relationships for the upper soil layer is given by Russo and Bresler (1981a). It is commonly assumed that K_s is the parameter of largest effect upon flow and transport. However, recently investigations of the spatial variability of retardation coefficients have been initiated (see Mackay et al. 1986a).

It is worthwhile to mention a few statistical moment properties of interest as revealed by field investigations:

(1) the degree of variability expressed by σ_Y^2 may be quite large, but in many cases σ_Y^2 was found to be smaller than unity;

(2) the correlation scale I_Y characterizing local heterogenity (Dagan 1986) is in the order of meters to tens of meters, and therefore much larger than the laboratory scale. In contrast, the log-transmissivity correlation length, characterizing spatial variability at the regional scale, is in the order of kilometers and

(3) for a limited set of measurements, the parameters of equation 5 and similar ones, can only be estimated, and these estimates are themselves subject to uncertainty. This is an additional source of uncertainty in prediction of flow and transport.

Modeling of Flow and Transport in Spatially Variable Formations

The basic equations 1-3 to determine θ, ψ, **q** and C apply at any point of the formation, but the soil properties reflected in the various coefficients are viewed as random functions of the space coordinate. Consequently, the dependent variables, solutions of these equations, are themselves random, and can be described in probabilistic terms only. Thus, rather than predicting the concentration as a deterministic function of **x** and t, we are able to derive only its expected value, variance, etc. This is a definite departure from the traditional approach and it has important implications on management.

The simplest, "brute force", approach to solving the flow and transport equations, is by Monte Carlo simulations (e.g. Smith and Schwartz 1980, for two-dimensional saturated flow and Andersson and Shapiro 1983, for one-dimensional unsaturated flow). The procedure consists of generating first realizations of the formation properties on a spatial grid, based on their known probability distribution functions. Thus, in the case of the lognormal K_s, the values of its logarithm constitute a multivariate normal vector, of given mean and covariance matrix, equation 5 (there are standard procedures to repetitively generate realizations of such a vector). Using a similar approach for other soil properties, equations 1-4 now represent, in each realization, flow and transport in a heterogeneous formation of given, deterministic, coefficients. These equations are subsequently solved, generally by numerical methods, to render the various dependent variables (in particular C) at the grid nodes and at different times. By repeating the procedure many times, a set of values is generated for C, which enable us to construct, empirically, its probability density function, i.e. its various statistical moments at each **x** and t. Although the scheme is conceptually simple, its implementation may require an enormous computational effort for three main reasons: (i) the solution of the transport problem in a heterogeneous given formation of a three-dimensional structure is mathematically complex and requires a large computer time, (ii) to obtain an accurate representation of the statistics of the dependent variables a large number of realizations must be generated, depending on the coefficients of variation of the input variables and on the estimation errors of parameters, and finally, (iii) to maintain the spatial correlation structure the grid has to be dense relative to the correlation scale, I_Y, and similar ones, while the flow domain is large compared to I_Y. Therefore, it seems, at present and in the near future, that fully blown Monte Carlo simulations will serve as numerical experiments rather than standard tools.

As an alternative, two paths can be followed in order to arrive at tractable methods to predict the statistical moments of concentration: (i) development of approximate, simplified, flow and transport models, and (ii) taking advantage of space averaging.

The first approach has been explored quite extensively in the emerging literature on stochastic modeling of subsurface flow. A few such approximations will be described in the third and fourth sections, and we recall two of the most widespread. Thus, in the case of unsaturated flow in the upper soil layer the flow is assumed to be vertical and heterogeneity manifests only in the variation of soil properties in the horizontal plane. This may be a valid approximation if the horizontal correlation scale is much larger than the depth of interest, say from the surface through the root zone. In groundwater flow a common approximation is a first-order expansion of the statistical moments of the dependent variables in the variance of the input parameter, e.g. in σ_Y^2, equation 5. This assumption simplifies considerably the computations; its main limitation is its applicability to $\sigma_Y^2 < 1$. Seeking simplified, approximate solutions of the flow and transport equations to facilitate the evaluation of the statistical moments is dictated not only by expediency.

A more profound reason is that statistical averaging has a smoothing effect, of error cancellation. Thus, even if results of prediction based on a simplified model are not accurate in each realization, the outcome for the expected value may be quite close to the exact one.

The second approach, that of exploiting the properties of space averaging, (which is of a more fundamental nature) is elaborated herein. The main point is that in many applications we are generally not interested in predicting the concentration at any point in the formation, but may be satisfied with its space average. In more concrete terms, let us consider first the example of leaching of a large agricultural field by uniform water application on the surface. With C a function of x,y,z and t, we concentrate on a horizontal plane at depth z. Due to spatial variability, C varies in a random fashion with x,y, but our interest lies primarily in its average value over the field, since the crop itself is distributed over it. Hence, the variable of interest is

$$\bar{C}(z,t) = \frac{1}{A} \iint_A C(x,y,z,t)\, dx\, dy \qquad [6]$$

where A is the domain in a horizontal plane at z. The space average $\bar{C}$ is a random variable and may be characterized by its statistical moments: the expected value, $\langle \bar{C} \rangle$, the variance, $\sigma^2_{\bar{C}}$, etc. These moments depend crucially on the ratio between L, the length scale of A, and I, the horizontal correlation scale, of the spatially variable properties. Now, we may take advantage of the fact that by their definition *nonpoint sources* are spread over large areas. Consequently, it is reasonable to assume that for such sources L/I>>1. A similar reasoning applies, for instance, to an aquifer discharging in a river: our main interest resides in the total quantity of solute entering the river, rather than in the local concentration at each point. This quantity is related to $\bar{C}$, equation 6, the averaging surface A being now a vertical plane along the boundary. Again, for nonpoint sources one may assume that L/I>>1.

The relationship between the statistical moments of $\bar{C}$ and those of C are discussed in detail by Vanmercke (1983) and in the present context by Dagan (1986). The main result concerns the variance, which by equation 6 is as follows

$$\sigma^2_{\bar{C}} = 1/A^2 \int_{A'} \int_{A} COV_C(x,x')\, dA\, dA' \qquad [7]$$

where COV_C is the two-point concentration covariance. Assuming that it is characterized by the same correlation scale, I, as the input variables, and for L>>I, equation 7 leads to the estimate

$$\sigma^2_{\bar{C}} \simeq \sigma^2_C \frac{I^2}{L^2} \qquad [8]$$

Hence, the important conclusion is that the variance of $\bar{C}$, related to spatial variability of soil properties, is very small if L>>I. Thus, $\bar{C}$ is approximately equal to its expected value, i.e. it is essentially a deterministic quantity (the same reasoning leads to the use of macroscopic variables at the laboratory scale and justifies the use of the representative elementary volume concept). Under these conditions we may ask the following questions: is it possible to derive $\bar{C}(z,t)$ as a solution of differential equations similar to equation 3, in which the actual soil properties are replaced by "effective" properties? If the answer is positive, how are these effective properties related to the statistics of the spatially variable properties?

There is not yet a definite answer for these questions in the case of unsaturated flow; some results will be discussed in the third section. If solute paths are statistically independent over most of A, it is reasonable to assume that equation 3 may be replaced by

$$R_{ef} \frac{\partial \bar{C}}{\partial t} + U \frac{\partial \bar{C}}{\partial z} = \frac{\partial}{\partial z} \left(D_{ef} \frac{\partial \bar{C}}{\partial z} \right) - S_{ef} \qquad [9]$$

where U = the average (arithmetic) vertical water velocity,
R_{ef} = the effective retardation,
S_{ef} = the irreversible sink strength, and
D_{ef} = the field-scale or "macrodispersion" longitudinal coefficient.

Although establishing the validity of equation 9 and relating the effective properties and parameters to their distributed values are still matters of debate and investigation, there is no doubt that equation 9 simplifies considerably the task of determining $\bar{C}$. On a firmer basis, discussed in a later section, it has been found that under certain conditions an equation similar to equation 9 is valid for aquifer flow, in which z is replaced by x, the horizontal direction of the mean flow. The main conclusions regarding the "macrodispersion" is that it is orders of magnitude larger than the pore scale dispersion, and that it is generally growing with the travel time or distance of the solute body, as will be shown.

Returning to equation 7, we have mentioned before that C is also subject to uncertainty due to errors of estimation of parameters, like those appearing in equation 5, or parameters pertaining to initial conditions, e.g. the solute concentration at the input zone. The variance of $\bar{C}$ associated with estimation errors is not diminished by the space-averaging process (see, for instance, Feinerman et al. 1986), so that the complete result is $\sigma^2_{\bar{C}} \simeq \sigma^2_C I^2/L^2 + \tilde{\sigma}^2_{\bar{C}}$, where the second term results from parameter estimation variances. This is the only source of uncertainty if $L>>I$, and its evaluation is a matter of further investigations.

Obviously, matters are more complicated in the case of point sources, like waste sites or repositories, for which the space average of the concentration over a large area does not characterize sufficiently the pollutant distribution, but this subject is beyond the scope of this review.

We shall touch again on a few points of principle in the final section, but first we review a few advances in stochastic modeling of transport.

SOLUTE TRANSPORT IN THE VADOSE (UNSATURATED) ZONE

The movement and retention of solutes in the vadose zone has always received much attention from soil scientists. Although the underlying physics is similar, groundwater transport and vadose zone transport do differ with respect to scale, flow regime, and the direction of the principle velocity components relative to the direction of the principle variations in the porous medium properties. Now we review two simplified approaches to model transport in heterogeneous soils, which have been developed in the last decade.

The Mechanistic Approach

The most common approach to analyze vadose-zone transport processes has been to model water and solute transport by using macroscopic quantities which vary in a deterministic manner, obey physical and chemical laws, and are expressed in the form of partial differential equations, e.g. the Richards' equation which combines equations 1 and 2, and the convection-dispersion equation (CDE) with linear equilibrium adsorption, equation 3.

The inherent spatial variability in the soil properties (e.g. Beckett and Webster 1971, Nielsen et al. 1973, Russo and Bresler 1981a, Jones and Wagenet 1984, among others) may limit the

applicability of the traditional deterministic approach, used to analyze water and solute transport processes, to field scale transport problems. Because of the stochastic nature of the transport parameters $K(\theta)$ and $\psi(\theta)$, and therefore of the water velocity $V = q/\theta$ = funct[K,ψ;BIC] and of the pore-scale dispersion coefficients $\underline{D}$=funct(λ,V), the soil system response (in terms of θ, ψ or C) to a given set of boundary and initial conditions (BIC) is also stochastic. This problem led to the development of stochastic models to describe water and solute transport (Dagan and Bresler 1979, Bresler and Dagan 1981, Amoozegar-Fard et al. 1982, Simmons 1982, Bresler and Dagan 1983), and water flow (Dagan and Bresler 1983, Andersson and Shapiro 1983, Yeh et al. 1985 and Mantoglou and Gelhar 1987).

The existing vadose-zone stochastic solute transport models differ in some of their underlying assumptions, boundary and initial conditions, and method of solution. They share, however, the common goal, namely, evaluation of the probability density function (PDF) of the solution concentration C(**x**,t), where **x** is the vector of the spatial coordinates.

Using the mechanistic approach for the field-scale solute transport, the problem can be formulated as follows: for a given field, whose spatial variability is characterized by the PDF of the $K(\theta)$ and the $\psi(\theta)$, and for given boundary and initial conditions, evaluate the PDF of C(**x**,t). In particular, prediction of the expected value of C and the relationships it satisfies, is of paramount importance.

A complete analysis of the transport of non-interacting solute through heterogeneous unsaturated soils under transient water flow conditions requires a knowledge of the functional relationships between K and θ and between ψ and θ as well as the knowledge of λ, and their joint PDF at various points throughout the system. Unfortunately, this information is generally not available, but in practice, the formidable task is greatly eased by employing several simplifying assumptions. A considerable simplification is introduced by restricting the quantification of a pertinent soil property, denoted generically by U(**x**), to its first and second statistical moments only. Furthermore, it is usually assumed that U(**x**) can be described by the general model

$$U(\mathbf{x}) = M(\mathbf{x}) + Z(\mathbf{x}) \qquad [10]$$

where $M(\mathbf{x})$ = a prior mean or drift function, and
$Z(\mathbf{x})$ = a zero-mean, stochastic stationary function characterized at second-order by the two-point covariance function.

In other words, according to equation 10, U(**x**) is characterized by variations at two different scales. The larger-scale variations are viewed as slowly-varying deterministic trends around which there are more localized variations which are viewed as locally stationary.

Strictly speaking, U(**x**) should be regarded as a three-dimensional, anisotropic stochastic function. The paucity of experimental data available from a given field site, however, usually limits the detection of spatial anisotropies. In addition, in most agricultural-oriented unsaturated-zone studies, the horizontal plane is much larger in scale than the vertical plane, and it is therefore common to neglect property variations with depth, and to describe U(**x**) at a given point in the horizontal plane by depth-averaged values, so that U(**x**) is interpreted as a two-dimensional isotropic stochastic function in the horizontal plane (Russo and Bresler 1981a).

To evaluate the response on the large field scale, based on scattered points or small-volume measurements, it is usually assumed that on the local scale, both equations 1 and 2 and also equation 3 apply. In addition, in order to simplify the stochastic analysis it is convenient to express the $K(\theta)$ and $\psi(\theta)$ functions in terms of analytical expressions characterized by a small number of parameters. Various parametric models have been used in the past. Thus, for $K(\theta)$, Nielsen et al. (1973), adopted a two-parameter exponential relationship. Russo and Bresler

(1981a,b) and Dagan and Bresler (1979, 1983) adopted the five-parameter presentation of Brooks and Corey (1964) for the $K(\theta)$ and the $\psi(\theta)$ functions. Andersson and Shapiro (1983) adopted the five-parameter presentation of van Genuchten (1980) for the $K(\theta)$ and the $\psi(\theta)$ functions. Yeh et al. (1985), and Mantoglou and Gelhar (1987) adopted the two-parameter exponential presentation of Gardner (1958) for the $K(\psi)$ function and assumed a linear relationship for the $\psi(\theta)$ function. In many of the stochastic approaches, however,(Dagan and Bresler 1979, 1983; Bresler and Dagan 1981, 1983a,b; Andersson and Shapiro 1983; Yeh et al. 1985), all parameters but the saturated hydraulic conductivity, K_s, were assumed to be deterministic constants, and the randomness of $K(\theta)$ or $K(\psi)$ stemmed only from the stochastic nature of K_s.

Initial investigations of unsaturated now and noninteracting solute transport in spatially-variable soils (Dagan and Bresler 1979, 1983; Bresler and Dagan 1981, 1983a,b) conceptualize the unsaturated flow system to be comprised of a collection of one-dimensional (vertical) independent soil columns without interaction between columns. The typical horizontal dimension scale of these hypothetical soil columns was not specified and was not needed for evaluation of the expected value of concentration. Each column, however, was assumed to be vertically homogeneous (i.e. possible property variations with depth were neglected and depth-averaged values were used), so that flow and transport in each of these columns can be described by equations 1-4, with S=0. Accordingly, the variability in the response of the field is the result of variations in the soil properties among the vertical soil columns.

To illustrate the stochastic mechanistic approach in the analysis of field-scale solute transport in the vadose-zone, we refer to the works of Dagan and Bresler (1979, 1983) and Bresler and Dagan (1979, 1981, 1983a,b). In addition to the above-mentioned assumption, they adopted the following: (i) the soil hydraulic properties are described by Brooks and Corey (1964) expression; (ii) the $\psi(\theta)$ relationships are deterministic, and the randomness of $K(\theta)$ stems from the stochastic nature of K_s through the scaling factor, δ

$$K_s(x)/K_s^* = \delta^2(x) \qquad [11]$$

where K_s^* is given by

$$K_s^* = \left[\int_0^\infty K_s^{1/2} f(K_s)\, dK_s \right]^2 = E(K_s) \exp(-2\mu_y - 2\sigma_y^2) \qquad [12]$$

where the scaling factor, δ, is assumed to be a second-order stationary lognormally distributed variate, i.e. $y+\ln\delta:N(\mu_y,\sigma_y)$, independent of its spatial position.

Additional assumptions for the simple case of steady gravitational water flow (Dagan and Bresler 1979, Bresler and Dagan 1981) are: (i) the flow is generated by a steady random recharge at a rate R, of a rectangular (uniform) distribution and (ii) water flow is steady, so that the pore water velocity and the soilwater content at any point in the horizontal plane of the field do not change with time and depth. For the case of transient water flow (Dagan and Bresler 1983) the assumptions were:

(1) unsteady water flow is generated by deterministic recharge, applied on the surface at a rate R during the infiltration time, t_i;

(2) water content is initially deterministic and uniform throughout the entire field;

(3) a piston flow-type profile is assumed at any point in the horizontal plane of the field during infiltration and subsequent redistribution.

Under these assumptions, at a given point in the horizontal plane of the field, the water flux q and the water content θ are step functions along vertical lines, with constant values at the point, for depths shallower than the position of the wetting front L(t) at time t, and with values equal to the initial ones, for depths beyond the wetting front. Accordingly, for the domain of $0<z<L(t)$, vertical pore water velocity $V=q/\theta$, depends only on time and can be approximated by $V(t)=dL(t)/dt$.

For transport they assumed an inert (conservative) solute, neglected the molecular diffusivity, and adopted equation 3 with S=0 for one-dimensional vertical flow, with longitudinal dispersivity coefficient λ_L being a lognormally distributed, second-order stationary variate, independent of δ. In the case of steady flow, under these simplifying assumptions, Bresler and Dagan (1981), provided closed-form expressions (requiring numerical quadrature at most) for various moments of the distribution of C(z,t). In the case of transient flow, they assumed that the dispersivity increases from zero to its assumed maximum value of 3 cm as the water front propagates downward. Assuming that the condition of zero solute gradient applies sufficiently far from z=L, Bresler and Dagan (1983b) provided closed form expressions for the first three moments of the distribution of C at different depths and times.

Although the results of the analyses of Dagan and Bresler (1979, 1983) and Bresler and Dagan (1981, 1983a,b) are based on simplifying assumptions, they may be applicable to the upper part of the vadose zone. The results suggest a few points of interest:

(1) during infiltration, the average concentration profile in a heterogeneous field is affected by the spatial variability of the hydraulic properties, by the average and the variability of the application rate, and (to a lesser extent, depending on the variance of y) by the average pore-scale longitudinal dispersivity coefficient. The spread of the average concentration profile generally increases with these factors;

(2) the average concentration profile in a heterogeneous field cannot generally be modeled as the solution of the classical convection-dispersion equation (CDE) equation 9 with constant effective coefficients. Thus in a heterogeneous field the width of the average concentration profile expands linearly with time, while in the solution of the CDE for an equivalent homogeneous fictitious soil, the transition zone expands as the square root of time, due to pore-scale dispersion. No attempt has been made, however, to model the mean concentration profile by an equation similar to equation 9 in which D_{ef} is *time dependent*. It is emphasized that one of the major sources of velocity variation is the ponding assumed to occur on that part of the field on which R exceeds K_s;

(3) nevertheless, under steady-state water flow, when the spatial variability in the soil hydraulic conductivity is relatively small, and when the average application rate is small compared with the mean saturated hydraulic conductivity, under moderate uncertainty in the application rate, the average concentration profile may be approximated by the solution of the CDE with constant effective coefficients;

(4) under transient water flow, the average concentration profile penetrates deeper than in steady flow. This is because during the transient period of buildup of the moisture content, the water content values are smaller than those of the ultimate steady state. Consequently, for a given recharge rate at the surface, the convective velocity is larger and the solute front propagates faster;

(5) under transient water flow conditions, an equivalent fictitious uniform porous medium cannot be defined simply by finding effective parameters for the governing PDE.

To demonstrate these points, figure 5 presents profiles of the mean salinity concentration for two different spatially-uniform recharge rates, and two different soils (highly-variable clay loam soil, and moderately-variable sandy loam soil), calculated by three approaches: (i) steady gravitational flow considering the spatial variability of the soil properties (Bresler and Dagan 1981); (ii) the same as (i) but under unsteady infiltration conditions (Bresler and Dagan 1983b), and (iii) steady gravitational flow in a fictitious equivalent uniform soil characterized by effective parameters (Bresler and Dagan 1983a). The calculated profiles in figure 5 clearly demonstrate the effects of both the soil hydraulic properties and the recharge rate on the shape of the mean salinity profiles. The deeper penetration of the mean salinity profile under transient flow conditions than under steady flow is also shown in figure 5. In addition, the failure of a deterministic approach (using effective, constant, parameters) to describe the mean salinity profile, particularly in a highly variable soil, is also shown.

In summary, the work of Dagan and Bresler (1979, 1983) and Bresler and Dagan (1981, 1983b) demonstrates that under field conditions, the dominant solute-spreading mechanism is spatial heterogeneity combined with convection. In other words, the spatial distribution of the pore water velocity, which, in turn, is controlled by the spatial distribution of the soil hydraulic properties and of the conditions at the soil surface, dominates the spatial distribution of the solute concentration, and consequently, the shape of the average concentration profile.

Non-mechanistic Approach

An alternative stochastic approach, non-mechanistic in form, is the transfer function model (TFM) applied to solute transport in soils by Jury (1982). This approach assumes that the internal physical mechanisms which contribute to the solute movement are unknown or cannot be known. Therefore, the soil water system is characterized entirely in terms of its ability to transform an input function (solutes added to the soil surface) into an output function (solutes moving through the soil). According to this perspective, the movement of solute through the unsaturated zone is governed by a travel-time PDF, $f_L(t)$, such that $f_L(t)dt$ is the probability that a solute entering the soil surface (z=0) at time zero will reach the depth z=L in the time interval from t to t+dt. By superposition, for an arbitrary application of solute at z=0, the concentration at z=L (Jury 1982) is

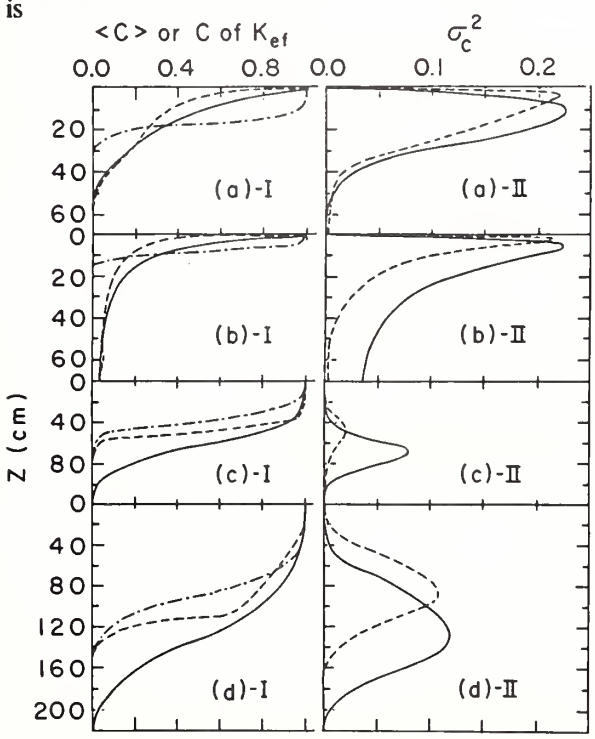


Figure 5.
Distribution with depth z during gravitational infiltration (dashed and dashed-dotted lines) and during unsteady infiltration (solid lines).(I) Dimensionless solute concentration C calculated with deterministic value of $K(\theta)$ equal to K_{ef} (dashed-dotted lines): expectation of C calculated according to Bresler and Dagan (1981), denoted by dashed lines and according to Bresler and Dagan (1983b) denoted by solid lines. (II) Variance of C with solid and dashed lines denoted as in (I). (a) Panoche soil, R=0.5 cm/h, t_i=24h. (b) Panoche soil, R=6.5 cm/h, t_i=6h. (c) Bet Dagan soil, R=0.5 cm/h, t_i=24h. (d) Bet Dagan soil, R=6.5 cm/h, t_i=6h.

$$C(L,t) = \int_0^{\infty} C_{in}(t-t')\, f_L(t')\, dt' \qquad [13]$$

where C_{in} is the concentration imposed at z=0.

Generally, $f_L(t)$ has to be determined by field measurements. To predict solute transport beyond the measurement zone (0<z<L), however, mechanistic assumptions are required. Under the assumption that convective vertical solute movement dominates lateral solute mixing, and that the profile is vertically homogeneous, the distribution of the physical processes contributing to the travel-time PDF, $f_L(t)$, between z=0 and z=L, is the same for all depths, and the travel-time PDF, $f_z(t)$, for any arbitrary depth, z, can be related to that for a reference depth L by the expression

$$f_z(t) = (L/z)\, f_L(tL/z) \qquad [14]$$

Moreover, in the case of steady-state water flow, both f_L and f_z can be translated from PDFs in terms of time to PDFs in terms of the depth of the net applied water. Hence solute transport becomes solely a function of the net amount of water applied, and not of the rate at which the water is added. Figure 6 shows calculated profiles of the mean salt concentration as a function of the net water applied, I, for an input pulse of $\Delta I=10$ cm, under a spatially uniform application rate. In this case variations in travel time from the soil surface to L=100 cm result from soil variations, and not from spatial variations of the application rate. Figure 2 illustrates the large spreading over depth which results from the travel-time (or the velocity) distribution. After the 10 cm pulse was moved into the soil by 50 cm of water, it dispersed over 300 cm and decreased to 15% of its initial value at the surface.

Jury et al. (1986) generalized the TFM of Jury (1982) to describe the movement of solute that may undergo physical, chemical or biological transformations as it moves through the soil, by introducing a solute lifetime PDF instead of the solute travel-time PDF. They showed that the generalized TFM is related to the law of mass balance for a solute as interpreted in the context of probability theory. They concluded that mechanistic stochastic models of solute movement, consistent with the law of mass balance, are also consistent with the TFM of Jury (1982).

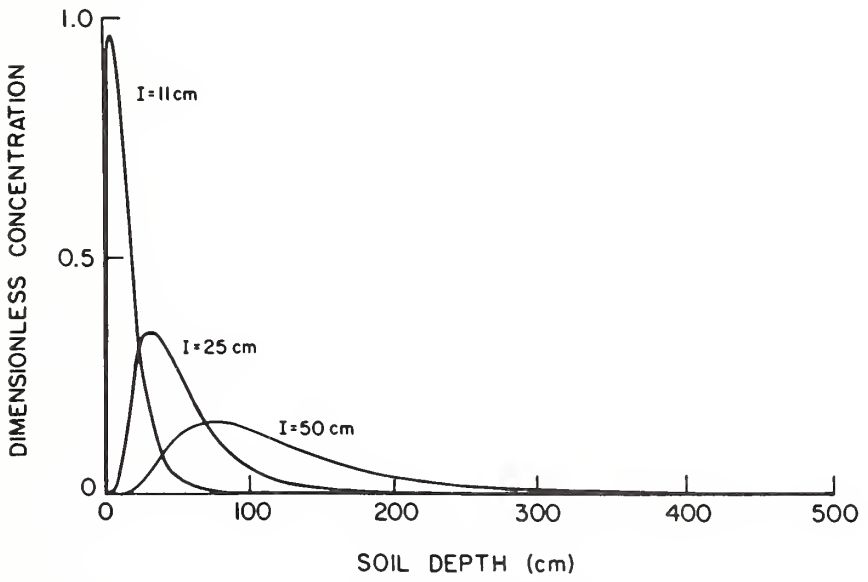


Figure 6.
Average solute concentration vs. depth for an inlet pulse width $\Delta I=10$ cm, for several values of I, the net water applied (from Jury 1982).

This can be also demonstrated by referring to the mechanistic stochastic model (MSM) of Dagan and Bresler (1979) which was discussed in the previous section. Based on field-scale solute concentration data, Jury (1982) suggested a parametric presentation for the PDFs f_L and f_z, using the two-parameter lognormal distribution function, characterized by a mean, μ, and variance, σ^2. If equation 14 is valid, these parameters are independent of depth. This means that the TFM predicts that the variance of the travel time is proportional to the square of the depth of observation, whereas the classical CDE predicts that the variance of the travel time is proportional to the depth itself. In other words, according to the TFM, the effective dispersivity will increase linearly with depth. This is consistent with the results of the MSM of Dagan and Bresler (1979), emphasizing the effect of the pore water velocity variations on solute spreading.

Discussion

The consistency between the results of the TFM of Jury (1982) and the MSM of Dagan and Bresler (1979) is not surprising since in both models the most important entities, the travel-time PDF $f_L(t)$ and the pore water velocity, $V=V(\delta,R)$, respectively, are approximately lognormally distributed. Both models neglect possible depth variation of their input variables, either δ or $f_L(t)$. In addition, both approaches focus on determining the expected value and variance of concentration, but not its two-point covariance in the plane or the moments of its space average equation 6. Under these conditions, and with neglect of lateral transport, knowledge of the possible correlation between nearby spatial points (characterized by a correlation scale) is not needed.

Another fundamental similarity between the two approaches is in the assumption that transport is dominated by vertical convection. Since the soil is regarded as a collection of homogeneous vertical columns, the displacement of a solute particle grows linearly with time in each column, and the same is true for any measure of the spread of the mean concentration profile. This corresponds to the Taylor (1921) short-time regime (see next section), for which the effective dispersion coefficient is time dependent. The picture is more complex, however, if pore-scale dispersion is accounted for, since it represents a mechanism by which displacements in the same profile become uncorrelated, but this is viewed as a minor effect compared to that of horizontal properties variability, if the latter is sufficiently large.

There are, however, differences between the TFM and MSM approaches. In the case of the TFM, for a given field site, the PDF of the travel time is assumed to be a function only of the net amount of water applied and not of the rate at which the water is added. In other words, the parameters describing $f_L(t)$ and $f_z(t)$ are independent of the surface boundary conditions. In the case of the MSM, however, the distribution of the pore water velocity, V, depends on both the spatial distribution of the inherent soil hydraulic properties (expressed in terms of the scaling factor, δ), and the conditions at the soil surface (expressed in terms of the distribution of the recharge rate, R). For a given field site, the parameters of the distribution of V are affected by both the mean and the variance of the recharge rate. It should be noted that if the variability of $V(\delta,R)$ is relatively small (due to relatively small variability in the soil properties, and relatively low average recharge rate), the PDF of V can be equally well approximated by a normal or lognormal distribution.

Inherent in the two approaches is the ensemble concept for defining the statistical properties of the input parameters and the output variable. Consequently, the ergodic hypothesis (e.g. Lumley and Panofsky 1964) must be invoked. This hypothesis states that inferences about the statistical structure of a stochastic function, U(x), may be based on estimates of the ensemble averages gained from spatial averages obtained from a single realization of U(x). For the spatial-averaging process to be meaningful, the spatial heterogeneities scale must be relatively small as compared with the overall scale of observation.

The analysis of Russo and Bresler (1982) suggested that in the case of a two-dimensional isotropic second-order stationary function, U(x), the error in estimating the ensemble averages of U(x), based on spatial averaging, depends on both the coefficient of variation, (CV), of U(x), and on the correlation scale, I_U, of U(x) relative to the length scale of the field. For example, for a soil property U(x), characterized by CV=0.7 and I_U=10 m, more than 100 measurements should be made over an area equivalent to more than 100 correlation scales, i.e. $A \simeq \pi(100I_U)^2 \simeq 300$ ha, in order to estimate the mean of U(x) with 20% error (relative deviation from the mean). These numbers, however may be reduced, if a larger error, say 40%, may be accepted. In this case, 30 measurements should be made over an area equivalent to 40 correlation scales, $A \simeq 50$ ha.

This example points to a serious difficulty associated with both the reliability of the models input parameters on one hand, and the interpretation and field validation of the models results, on the other hand. For example, the PDF of the travel time $f_L(t)$, of the TFM of Jury (1982) was estimated from the average of 14 measurements of bromide concentration taken from a depth of 30 cm of 14 sites, uniformly distributed over a 1.5-ha field (Jury et al. 1982). Estimates of the field-average bromide concentration based on the same 14 spatial points at other soil depths (60, 90, 120, and 180 cm), were then used to validate the TFM as calibrated previously from the 30-cm depth bromide data. The statistical significance of the calibrated $f_L(t)$ and of the sample averages of the bromide concentrations, viewed as representing ensemble averages, is problematic.

In the case of the MSM of Dagan and Bresler (1979) or that of Bresler and Dagan (1981), estimates of the PDF of δ were based on 120 sample values obtained from 20 sites randomly distributed over a 150-ha Panoche field, (Warrick et al. 1977), or on 120 values obtained from 30 sites randomly distributed over a 0.8 ha-Bet Dagan field, (Russo and Bresler 1980). Based on the analysis of Jury et al. (1987), who found for both fields a correlation length of $I_U \simeq$ O.7m for δ derived from measurements of $K(\theta)$, we can conclude that for both fields, the error associated with estimates of the PDF of δ is not too large. A difficulty arises, however, when an attempt to validate the MSM results is made, based on a relatively small sample size. We demonstrate this point with data taken from a recent field study, in which the field-average chloride concentration data were estimated using large-diameter undisturbed soil cores from 16 spatial locations randomly distributed in the Bet Dagan field. To evaluate the error, ϵ, associated with the estimates of the ensemble mean of the chloride concentration, we used data together with calculated results of Russo and Bresler (1981b) for the relationships between the correlation scale of the solute concentration and the average solute concentration during leaching, and we applied the approach of Russo and Bresler (1982). In figure 7, the sample mean chloride concentration and the interval, i.e. mean $(1\pm\epsilon)$ are plotted as a function of time. The error is relatively large at the initial stages of the leaching process (ϵ=90% of the mean, for t=1 hr) but decreases with time (ϵ=20% of the mean at t=4 hr) as a larger portion of the field is completely leached. It should be noted, however, that the decrease in ϵ with time (fig. 3) is specific to the case of leaching where a solute concentration of C_0 is continuously added at the soil surface. Consequently, for a given soil depth, a larger portion of the field is leached, and the spatial variability of the solute concentration decreases, as time increases. Mean chloride concentration calculated by the stochastic approach of Dagan and Bresler (1979) disregarding pore-scale dispersivity, i.e. λ=0, by the stochastic approach of Bresler and Dagan (1981) (using a deterministic λ=3 cm) and by a deterministic approach, are also shown in figure 7. It is seen that the sample mean values are fitted best by the mean concentration values calculated by the stochastic approach of Bresler and Dagan (1981), but the mean concentrations calculated by the three different approaches are within the aforementioned error interval.

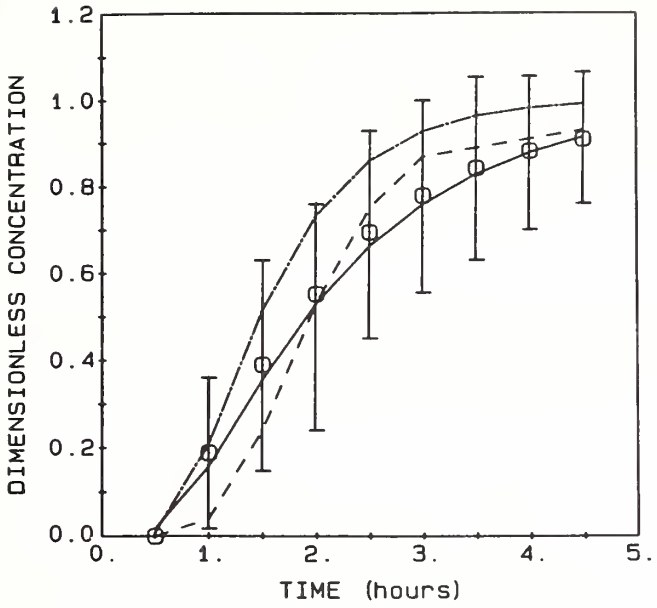


Figure 7.
Mean chloride concentration as a function of time. Sample means calculated from experimental data from 16 locations in Bet Dagan field are given by the blank squares. The small horizontal bars indicate the interval (expected value ± ϵ), where the error ϵ, is calculated from Russo and Bresler (1982). Dashed and solid lines represent mean values computed from Dagan and Bresler (1979) for piston flow, and Bresler and Dagan (1981) for dispersive flow, respectively. Dashed-dotted line represent deterministic calculation of the chloride concentration, using deterministic (average) values of V and λ=D/V.

The one-dimensional models, the MSM of Dagan and Bresler (1979) and the TFM of Jury (1982), neglect transport in the other two dimensions. This neglect stems from their concern with transport in the soil upper layer and the assumption that its depth is much smaller than the horizontal correlation scale of soil properties affecting transport. However, if transport takes place to relatively large depths (e.g. to a deep water table) the heterogeneity in the vertical direction may influence the process. It is important in such cases to evaluate the effects of three-dimensionality of the property variations, on transport. The problem is extremely complex and so far transport has not been studied in a theoretical frame. There are, however, a few investigations of the flow, and we refer to the work of Yeh et al. (1985), who analyzed steady unsaturated water flow with vertical mean infiltration through unbounded heterogeneous porous media, using a first-order perturbation approximation in σ_Y^2. They assumed that the unsaturated hydraulic conductivity is related to the soil water pressure head by the two-parameter exponential model (Gardner 1958), and that the saturated conductivity, K_s, and the soil constant, α, are related to the pore-size distribution. In their approach, α was assumed to be a deterministic constant whereas K_s variability was represented by a three-dimensional, isotropic, and stationary random field characterized by an isotropic exponential covariance function. Using spectral analysis techniques, they calculated variances and covariance functions of the pressure head, effective hydraulic conductivities, and variances of the unsaturated conductivity, the pressure gradient, and the water flux; all as a function of the variance of $Y=\ell n\ K_s$, of its correlation scale, I_Y, and of deterministic values of α. We shall recall here some of their results concerning the specific discharge variance.

The analysis (Yeh et al. 1985) of three-dimensional flows and mean unit hydraulic gradient, suggested that the normalized longitudinal flux variance, $\sigma_q^2/[K(\theta)]^2\sigma_Y^2$, is constant, independent of the product αI_Y, as long as $\alpha I_Y \leq 1$, and decreases with αI_Y when $\alpha I_Y \geq 1$. The rate of decrease, however, is larger for increasing α. This type of dependence on the two parameters is shown in figure 8, in which the coefficient of variation of q, CV_q, is represented as function of I_Y for a few values of α. The correlation scale I_Y of $Y=\ell n\ K_s$, can be interpreted as a characteristic length of the variability of Y and α^{-1} can be interpreted as the height of the capillary fringe, or as a characteristic length of the flow domain above the saturated zone (water table). Hence, a small α is related to a fine-textured soil, whereas a large α pertains to a coarse-grain material. Since in saturated flow σ_q is independent of I_Y, under the same assumptions and flow conditions, it is seen that the mechanism prevailing in the unsaturated regime, represented in this analysis by the

parameter α, causes a reduction of the specific discharge variability. It seems, therefore, that the variation in the water content θ tends to reduce the variability of K, which controls primarily the variability of q. This reduction is larger for coarse-textured soils and large correlation scales.

In the next section we discuss a few points of principle about the dispersion effect of heterogeneity, which is proportional to σ_q^2 and to I_Y, if *transport takes place over a travel distance which is large compared to* I_Y. Hence, the results of Yeh et al. (1985) suggest that the effect of the flow three-dimensionality is to reduce the flux variability, and consequently the dispersive effect of heterogeneity, as compared to predictions based on simplified, one-dimensional models. The overall impact of three-dimensional structures on transport in unsaturated soil is a subject awaiting further investigations.

The relatively small variability in pore-water velocity, V, in coarse-textured soils, and its impact on the solute spread is demonstrated by the data of Schulin et al. (1987), obtained from a tracer experiment in a 15m long, 3m deep transect, under natural soil, vegetative, and climatic conditions (fig. 3). Analysis of the data showed a relatively uniform displacement in the vertical direction, as well as a significant horizontal redistribution during the study periods (200 days and 400 days, for chloride and bromide, respectively). Analysis of their 57 bromide and 61 chloride local concentration profiles (assuming local one-dimensional vertical steady flow, and the applicability of the classical CDE) revealed a relatively small variability of V, (CV<25%) which can equally be described by either a normal or a lognormal distribution function. Consequently, the classical CDE model and a stochastic model equally describe momentary field-averaged concentration profiles. A comparison of average (of local values) dispersivity, λ=3.7, and 2.1 cm, for the bromide and the chloride tracers, respectively, with the respective effective field scale dispersitivities, λ_{ef}=6.8 and 8.4cm, however, suggests that for both tracers, an equivalent fictitious homogeneous porous medium cannot be defined simply by finding mean values of the parameters of the CDE.

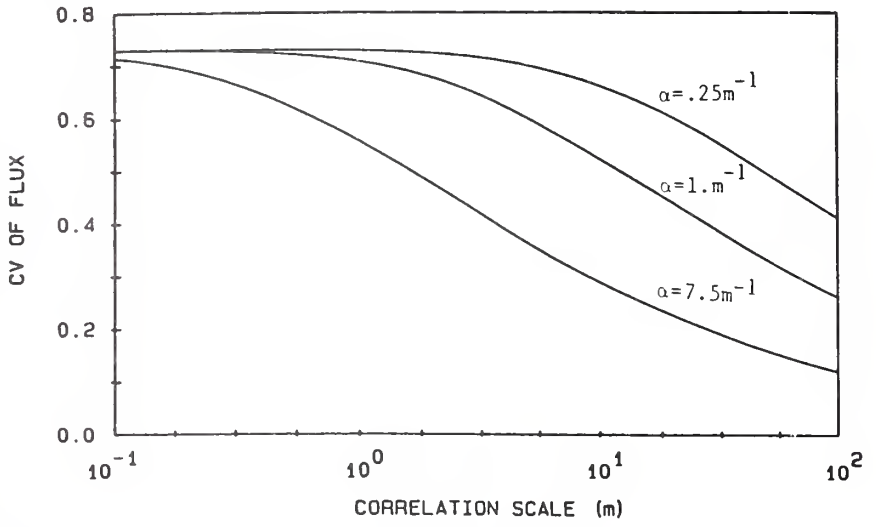


Figure 8.
Coefficient of variation of the longitudinal (vertical) flux CV_q, for three-dimensional flow in isotropic heterogeneous soil, as a function of the correlation scale I_Y of $Y = \ell n\, K_s$, for three different values of the soil constant α. CV_q is calculated from equation [47] of Yeh et al. (1985a), for mean unit vertical hydraulic gradient.

TRANSPORT BY GROUNDWATER

Transport in the vadose zone is generally caused by nonpoint agricultural sources. In contrast, solute movement by groundwater is related to a variety of sources, point or nonpoint. The flow mechanism is much simpler in the saturated medium than in the unsaturated one, since the hydraulic properties and the effective porosity are not influenced by the flow regime. However, it is essential to account for the three-dimensional nature of heterogeneity and flow, since solute generally travels over many heterogeneity correlation scales. As in the previous sections, we shall refer here mainly to nonreactive, neutrally buoyant solutes. The review here will be concise, since the field has been covered recently by a few surveys (Dagan 1986, 1987; Gelhar 1986).

The phenomenon of transport in porous media has been initially explored under laboratory conditions by experiments in porous columns. The traditional approach was to model transport with the aid of equation 3, in which $\underline{D}$ stands for the pore-scale dispersion tensor. For the usual large values of the Peclet number; Ud/D_d, where D_d is the coefficient of molecular diffusion, U is the average pore-scale velocity and d is the pore size; $\underline{D}$ is proportional to U in the λ_L (longitudinal) and λ_T (transverse) directions. Laboratory experiments have led to values of λ_L of the order of millimeters, while λ_T is smaller by an order of magnitude.

Subsequent field experiments have shown, however, that in natural formations the apparent dispersivities are larger than laboratory values by a few orders of magnitude (a compendium of field results is presented by Gelhar 1986). Furthermore, unlike the laboratory results, they tend to increase with the distance travelled by the solute from its source. These differences have been attributed to the effect of large-scale heterogeneity which is omnipresent, albeit at various extents, in different aquifers (Freeze and Cherry 1979). It is only in the last decade, however, that the phenomenon has been studied systematically in the frame of the emerging stochastic theory of groundwater flow (see Dagan 1986, for a review). It is accepted nowadays that:

(1) the spread of solutes is dominated by the large scale spatial variations of K, which changes irregularly in space (see fig. 4 for illustration);

(2) as a consequence, C, the solute concentration varies also in an irregular fashion (figs. 1,2);

(3) these variations are subject to uncertainty and the proper mathematical setting is the stochastic one, as discussed in the Introduction;

(4) in the case of nonpoint sources, i.e. for solute bodies or plumes whose initial dimensions are *large compared to the heterogeneity scale*, the expected value of the space-averaged concentration $\bar{C}$ (see equation [6]) satisfies a CDE similar to equation 9, i.e.

$$\frac{\partial \bar{C}}{\partial t} + \mathbf{U}\nabla\bar{C} = \nabla(\underline{D}_{ef}\,\nabla C) \qquad [15]$$

where $\mathbf{U} = \langle \mathbf{V} \rangle$, $\mathbf{V} = \mathbf{q}/\theta_s$, is the average horizontal groundwater velocity, and $\underline{D}_{ef}$ is the effective, field-scale, dispersion tensor.

Equation 15 has to be supplemented by retardation and irreversible sink terms in the case of reactive solutes (see equation 9). The crux of the matter is to determine the dependence of $\underline{D}$ upon the velocity and the heterogeneous structure, and this is the object of the following discussion.

When dealing with flow or transport in a heterogeneous formations, it is essential to distinguish between the different scales affecting the process. The relevant scales (Dagan 1986) are:

(1) the extent of the flow domain, i.e. the aquifer thickness, D, and its horizontal dimension, L;

(2) the dimension of the initial solute body or plume, of volume V_0;

(3) the extent of the area, A, over which $\bar{C}$ (equation 6) is averaged; e.g. the outlet of the aquifer to a river;

(4) the average distance Ut traveled by the solute body from the source, where t is the travel time and

(5) heterogeneity scales, i.e. the correlation scales of $Y = \ell nK$, I_{Yh} (horizontal) and I_{Yv} (vertical).

Among these, I_Y plays a key role and two main categories of heterogeneity have been suggested, according to its magnitude (Dagan 1986): the local and the regional. The local scale refers to the variability of K as measured by cores or other devices (e.g. slug test) which are much larger than the pore-scale, but small compared to D. An example of variability at this scale is presented in figure 4, for which it was found (Sudicky 1986) that $I_{Yv} \simeq 0.1$ m and $I_{Yh} \simeq 2.8$ m. The regional scale is the one characterizing the variability in the horizontal plane of the transmissivity, T, which is a space average of K over the entire thickness and a comparable planar dimension. Obviously, I_Y, where hereafter $Y = \ell nT$, is, in this case, much larger than D, but smaller than L. Comprehensive analyses of field data by Hoeksema and Kitanidis (1984) indicate that T varies irregularly in the plane and I_Y is of the order 10^3-10^4 m. The effects of these two types of heterogeneity upon transport are quite different and will be discussed separately.

The Local Scale

The local scale heterogeneity (fig. 4) causes enhanced spreading of the solute, due to the tortuosity of the paths of various solute parcels, but mainly because of the velocity variations along these paths. In the case of nonpoint sources the basic inequality is $V_0 >> I_Y^3$ and $A >> I_Y^2$. As a consequence, although the point value of C might vary considerably in space, C is a smooth function of x and t and the ergodic hypothesis is bound to be obeyed. In simple words, $\bar{C}$ in the actual realization is equal to the ensemble average $\langle \bar{C} \rangle$ (see Introduction), while its variance is close to zero. Then, prediction of the solute fate reduces to the determination of $\bar{C}$, and mainly to the effective dispersion coefficients, by using models of flow and transport.

The simplest models of heterogeneous formations (e.g. Mercado 1967, Marle et al. 1967, Matheron and De Marsily 1980, Güven et al. 1986) are of perfectly-layered aquifers (stratification is common in sedimentary formations). With z a vertical coordinate, the hydraulic conductivity, K, is a function of z solely. Since K can be seldom mapped accurately and is subject to uncertainty, it is regarded as a random function, stationary in its simplest representation. If K is lognormal, $Y = \ell nK$ is characterized by its mean, m_Y, and covariance, C_Y, equation 5, the only integral scale being the vertical one. Let us assume now that a uniform and horizontal water head gradient, J, drives the flow. By Darcy's Law, equation 2, the velocity is given by $V = q/\theta_s = KJ/\theta_s$, and it is also a random function of z, of mean $U = K_A J/\theta_s$, where K_A is the conductivity arithmetic mean. Assume now that a solute particle is inserted in the formation, and neglect the effect of pore-scale dispersion. Convection by water results in a horizontal motion of the particle with constant velocity V, pertaining to the layer at elevation z. Under this simple condition, it is easy to

calculate the average displacement of a solute body which extends over many layers, as well as its longitudinal spread (the "moment of inertia"). If the latter is regarded as the result of an effective dispersion process, the closed-form expression of D_L, the longitudinal coefficient, is as follows

$$D_L = \frac{\sigma_K^2}{K_A^2} U^2 t \; ; \; \lambda_L = \frac{D_L}{U} = \frac{\sigma_K^2}{K_A^2} Ut \qquad [16]$$

This simple and fundamental result has a few interesting implications. First, it is similar to the early-time regime of diffusion by continuous motions derived by Taylor (1921) for any process for which the travel time is small compared to that of lateral mixing by pore-scale dispersion. Note also that the scheme leading to equation 16 is similar to the one employed by Dagan and Bresler (1979) to model vertical transport in the upper unsaturated zone, discussed in the previous section. The two striking features of the effective dispersivity λ_L, equation 16, are: (a) its value may be much larger than the pore-scale dispersivity and (b) it increases linearly with the travel time or distance. These features may explain in principle the field findings mentioned before. However, this model is of limited *quantitative* validity, except for transport over short distances. Indeed, its main weakness is the assumption of perfect layering, i.e. of homogeneity in the horizontal direction over the large distances travelled by the solute. It seems that in nature Y also varies with the horizontal coordinates x and y, though the correlation scale, I_{Yh}, may be much larger than I_{Yv}, as it is the case for the Borden site (fig. 4). Hence, a realistic modeling of transport must account for the three-dimensional heterogeneous structure, i.e. to regard C_Y, equation 5, as a function of x,y,z, generally with different correlation scales. Consequently, the solute is convected by V, whose *three* components are random, even if its mean U is constant. The derivation of the effective dispersivity under these conditions is a difficult task. Besides the numerical work of Smith and Schwartz (1981), limited to two-dimensional flow and particular cases; results of a general nature have been obtained only under the assumption that σ_Y^2 is sufficiently small, say $\sigma_Y^2 < 1$. The main contributions in this respect are those of Gelhar and Axness (1983), Dagan (1982, 1984, 1988) and Neuman et al. (1987). The main results of these studies can be summarized as follows:

(1) Under transport with average groundwater velocity U, the longitudinal dispersion associated with convection and heterogeneity is much larger than that related to pore-scale, provided that I_Y is much larger than the pore scale, which is generally the case.

(2) λ_L grows with travel time at the beginning of the solute motion, as predicted by equation 16, but tends to a constant effective dispersivity after a travel distance of a few tens of integral scales, I_{Yh}. Dagan (1982, 1984, 1988) has provided closed-form expressions for λ_L as function of time, which compared favorably with field measurements at the Borden site experiment (Freyberg 1986, Sposito and Barry 1987).

(3) The asymptotic value of λ_L, common to most above studies, is $\lambda_L \rightarrow \sigma_Y^2 I_{YL}$, where I_{YL} is the integral scale of Y along a line parallel to the direction of U, the average velocity.

(4) Matters are more complex with regard to the transverse dispersion.

By solving the time dependent evolution of the "moment of inertia" of the solute body by convection only, Dagan (1982, 1984) has arrived at the conclusion that it first grows with time, but it tends asymptotically to a constant value, while the associated λ_T first grows and then decreases. This result is in qualitative agreement with observed plumes which display little lateral expansion, and this picture is in agreement with the aforementioned field results. Consequently, Gelhar and Axness (1983) and Neuman et al. (1987), who sought from the outset the asymptotic value of λ_T,

could arrive at a finite value only by accounting for pore-scale dispersion, and the value is very small.

Very little is presently known about the spatial variability of the properties characterizing exchange processes between solutes and matrix, e.g. the retardation coefficient (see equation 9). Preliminary results from the Borden site (Mackay et al. 1986a) indicate that the retardation coefficient is spatially variable, but could not be correlated with other properties.

The Regional Scale

The heterogeneity at the regional scale is characterized by the transmissivity correlation scale, I_Y, which is much larger than the aquifer thickness, being of the order 10^3-10^4 m. The basic question is whether for nonpoint sources one is entitled to assume that the planar dimension of the initial solute body or plume, A_0, is much larger than I_Y^2. If this is the case, the same properties mentioned before are obeyed by $\bar{C}$, space averaged over an area or line of much larger extent than I_Y. Then, the effective dispersivity grows with time and tends to the asymptotic limit mentioned before only after a very large travel distance of the order of tens I_Y. In contrast, if A_0 is not large with respect to I_Y^2, the transport process has a completely different nature. The transmissivity variations do not result in an increased effective dispersivity, but in a winding motion in the plane, while the solute plane preserves its identity. In the stochastic context this manifests in lack of ergodicity, i.e. the concentration in the actual formation is very different from the ensemble mean. The diagnosis of this effect is found by computing the concentration variance, which may be quite large. These theoretical aspects have been analyzed recently by Dagan (1984), but have not yet been corroborated by field measurements at the regional scale. It was suggested that the approach to reduce the large uncertainty of predicting $\bar{C}$ prevailing at the regional scale, is to use conditional probability or geostatistical methods in representing the spatial variation of T (Dagan 1984). This topic, however, is beyond the scope of the present review.

SUMMARY AND CONCLUSIONS

The main message in this review is that heterogenity, i.e. spatial variability of soil macroscopic properties causes the enhanced spreading of solutes under field conditions, as compared with transport in laboratory columns. Furthermore, due to the irregular spatial variations and to the practical impossibility to measure and to map these properties in detail, they are subject to uncertainty. The same is true for dependent variables like water head, velocity or solute concentration. The stochastic models of transport, developed mainly in the last decade, treat properties and variables as random space functions. These models predict the expected value $\langle C \rangle$ and variance σ_C^2 of the concentration, as functions of space and time, rather than the traditional deterministic C.

Nonpoint sources are characterized by the large size of the input zone as compared to the heterogeneity scale. Furthermore, in applications one is often satisfied with predicting the space-averaged concentration, $\bar{C}$, over a large outlet area (e.g. a horizontal plane at some depth beneath the soil surface for the unsaturated zone or the boundary of an aquifer with a river). Under these conditions spatial variability does not cause appreciable uncertainty of $\bar{C}$ and prediction may be limited to its expected value. This might be true for transport in the unsaturated zone or at aquifer local scale, but may not be valid for regional scale heterogeneity. In the latter case, the variance of $\bar{C}$ may be quite large, unless geostatistical methods are employed in order to reduce it.

Besides spatial variability, the predicted concentration is subject to uncertainty due to errors of estimation of various parameters making up the models. This parametric uncertainty can be treated by classical statistical methods.

Existing models of transport in spatially variable formations are of a simplified nature. Thus, in the unsaturated zone, they are mostly based on the assumption of a vertical, one-dimensional flow. Still, their application requires a large body of field data concerning the variation of soil properties in the horizontal plane. The applicability of more sophisticated models, of a three-dimensional nature, does not seem warranted at present because of scarcity of adequate field data. There is an immediate need for large scale, controlled field experiments, in order to validate existing models. In the case of aquifer flow, existing simplified models account for the effect of local scale spatial variability upon the effective dispersivity. Less is known about the effect of regional scale heterogeneity upon transport. In all cases, the information about field variability of the properties affecting behavior of reactive solutes, is meager.

On the conceptual level, it is suggested that the main effect of the development of recent flow and transport models on management is to adopt a probabilistic, risk-based approach, rather than the usual deterministic, one. Then, the uncertainties caused by spatial variability and estimation errors, are set in a proper perspective, in terms of their influence upon the decision-making process.

REFERENCES

Amoozegar-Fard, A., D.R. Nielsen, and A.W. Warrick. 1982. Soil solute concentration distributions for spatially varying pore water velocities and apparent diffusion coefficients. Soil Sci. Soc. Am. J. 46:3-8.

Andersson, J., and A.M. Shapiro. 1983. Stochastic analysis of one-dimensional steady-state unsaturated flow: A comparison of Monte Carlo and perturbation methods. Water Resour. Res. 19:121-133.

Beckett, P.H.T., and R. Webster. 1971. Soil variability. A review Soils Fert. 34:1-15.

Bresler, E., and G. Dagan. 1979. Solute dispersion in unsaturated heterogeneous soil at field scale. 2. Applications. Soil Sci. Soc. Am. J. 43:467-472.

Bresler, E., and G. Dagan. 1981. Convective and pore scale dispersive solute transport in unsaturated heterogeneous fields. Water Resour. Res. 17:1683-1693.

Bresler, E., B.L. McNeal and D.L. Carter. 1982. Saline and sodic soils. Springer Verlag, Heidelberg, Berlin, and New York.

Bresler, E., and G. Dagan. 1983a. Unsaturated flow in spatially variable fields. 2. Application of water flow models to various fields. Water Resour. Res. 19:421-428.

Bresler, E., and G. Dagan. 1983b. Unsaturated flow in spatially variable fields. 3. Solute transport models and their application to two fields. Water Resour. Res. 19:429-435.

Brooks, R.H., and A.T. Corey. 1964. Hydraulic properties of porous media. Hydrol. Pap. 3, 27 pp., Colo. State Univ., Fort Collins.

Dagan, G. 1982. Stochastic modeling of groundwater now by unconditional and conditional probabilities. 2. The solute transport. Water Resour. Res. 18(4):835-848.

Dagan, G. 1984. Solute transport in heterogeneous porous formations. J. Fluid Mech. 145:151-177.

Dagan, G. 1986. Statistical theory of groundwater flow and transport: pore to laboratory, laboratory to formation and formation to regional scale. Water Resour. Res. 22:120S-135S.

Dagan, G. 1987. Theory of solute transport by groundwater. Ann. Rev. Fluid Mech. 19:183-215.

Dagan, G. 1988. Time-dependent macrodispersion for solute transport in anisotropic heterogeneous aquifers. Water Resour. Res., in press.

Dagan, G., and E. Bresler. 1979. Solute dispersion in unsaturated heterogeneous soil at field scale. I. Theory. Soil Sci. Soc. Am. J. 43:461-467.

Dagan, G., and E. Bresler. 1983. Unsaturated flow in spatially variable fields. 1. Derivation of models of infiltration and redistribution. Water Resour. Res. 191:413-420.

Feinerman, E., G. Dagan, and E. Bresler. 1986. Statistical inference of spatial random functions. Water Resour. Res. 22(6).

Freeze, R.A. 1975. A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media. Water Resour. Res. 11:725-741.

Freeze, R.A. and J. Cherry. 1979. Groundwater, Prentice Hall.

Freyberg, D.L. 1986. A natural gradient experiment on solute transport in a sand aquifer. 2. Spatial moments and the advection and dispersion of nonreactive tracers. Water Resour. Res. 22:2031-2046.

Gardner, W.R. 1958. Some steady state solutions of unsaturated moisture flow equations with application to evaporation from a water table. Soil Sci. 85:228-232.

Gelhar, L.J., and C.L. Axness. 1983. Three-dimensional stochastic analysis of macrodispersion in aquifers. Water Resour. Res. 19:161-180.

Gelhar, L.J. 1986. Stochastic subsurface hydrology from theory to applications. 1986. Water Resour. Res. 22:135S-145S.

Güven, 0. and F.J. Molz. 1986. Deterministic and stochastic analyses of dispersion in an unbounded stratified porous medium. Water Resour. Res. 22:1565-1574.

Hoeksema, R.J., and P.K. Kitanidis. 1984. An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling. Water Resour. Res. 20:1003-1020.

Jones, A.J., and R.I. Wagenet. 1984. In-situ estimation of hydraulic conductivity using simplified methods. Water Resour. Res. 20:1620-1626.

Jury, W.A. 1982. Simulation of solute transport with a transfer function model. Water Resour. Res. 18:363-368.

Jury, W.A., L.H. Stolzy, and P. Shouse. 1982. A field test of the transfer function model for predicting solute transport. Water Resour. Res. 18:369-375.

Jury, W.A., G. Sposito, and R.E. White. 1986. A transfer function model of solute transport through soil. 1. Fundamental concepts. Water Resour. Res. 22:243-247.

Jury, W.A., D. Russo and G. Sposito. 1987. The spatial variability of water and solute transport properties in unsaturated soil: II. Scaling models of water transport. Hilgardia 55:33- 56.

Lumley, J.L., and A. Panofsky. 1964. The structure of atmospheric turbulence. John Wiley & Sons, New York.

Mackay, D.M., W.P. Ball, and M.G. Durant. 1986a. Variability of aquifer sorption properties in a field experiment on groundwater transport of organic solutes: methods and preliminary results. Jour. Contam. Hydrol. 1:119-132.

Mackay, D.M., D.L. Freyberg, P.V. Roberts, and J.A. Cherry. 1986b. A natural gradient experiment in a sand aquifer. 1. Approach and overview of plume movement. Water Resour. Res. 22:2017-2030.

Mantoglou, A., and L.W. Gelhar. 1987. Stochastic modeling of large-scale transient unsaturated flow systems. Water Resour. Res. 23:37-46.

Marle, C., Simandoux, P. Pacsirsky, J.C. Gaulier. 1967. Etude du displacement de fludes miscibles en milieu poreux stratifie,. Rev. Inst. Fr. Pet. 22:272-94.

Matheron, G., and G. de Marsily. 1980. Is transport in porous media always diffusive? A counterexample. Water Resour. Res. 16:901-917.

Mercado, A. 1967. The spreading pattern of injected waters in a permeability stratified aquifer. IAHS/AISH Publ. No. 72, 23-36.

Moltyaner, G.L. 1987. Mixing cup and through-the-wall measurements in field-scale tracer tests and their related scales of averaging. J. Hydrol. 89:281-302.

Neuman, S.P. 1982. Statistical characterization of aquifer heterogeneities: an overview. Recent Trends in Hydrogeology, Geol. Soc. Am. Boulder, Colo., 81-102.

Neuman, S.P., C.L. Winter, and C.N. Newman. 1987. Stochastic theory of field-scale Fickian dispersion in anisotropic porous media. Water Resour. Res. 23:453-466.

Nielsen, D.R., J.W. Biggar, and K.T. Erh. 1973. Spatial variability of field-measured soil-water properties. Hilgardia. 42:215-260.

Nielsen, D.R., M.Th. van Genuchten, and I.W. Biggar. 1986. Water flow and solute transport processes in the unsaturated zone. Water Resour. Res. 22:89S-108S.

Russo, D., and E. Bresler. 1980. Scaling soil hydraulic properties of a heterogeneous field. Soil Sci. Soc. Am. J. 44:681-683.

Russo, D., and E. Bresler. 1981a. Soil hydraulic properties as a stochastic processes. 1. An analysis of field spatial variability. Soil Sci. Soc. Am. J. 45:682-687.

Russo, D., and E. Bresler. 1981b. Effect of field variability in soil hydraulic properties on unsaturated water and salt flows. Soil Sci. Soc. Am. J. 45:675-681.

Russo, D., and E. Bresler. 1982. Soil hydraulic properties as a stochastic processes: 11. Errors of estimates in a heterogeneous field. Soil Sci. Soc. Am. J. 46:20-26.

Schulin, R., M.Th. van Genuchten, H. Flühler, and P. Ferlin. 1987. An experimental study of solute transport in a stony field soil. Water Resour. Res. 23:1785- 1794.

Simmons, C.S. 1982. A stochastic-convective transport representation of dispersion in one-dimensional porous media systems. Water Resour. Res. 18:1193-1214.

Smith, L., and R.A. Freeze. 1979. Stochastic analysis of steady state groundwater flow in bounded domain. 1. One-dimensional simulations. Water Resour. Res. 15:521-528.

Smith, L., and F.W. Schwartz. 1980. Mass transport. 1. Stochastic analysis of macrodispersion. Water Resour. Res. 16:303-313.

Sposito, G.W., and D.A. Barry. 1987. On the Dagan model of solute transport: foundational aspects. Water Resour. Res. 23:1867-1875.

Sposito, G.W., W.A. Jury, and V.K. Gupta. 1986. Fundamental problems in the stochastic convection-dispersion model of solute transport in aquifers and field soils. Water Resour. Res. 22:77-88.

Sudicky, E.A. 1986. A natural-gradient experiment on solute transport in a sand aquifer: spatial variability of hydraulic conductivity and its role in the dispersion process. Water Resour. Res. 22:2069-2082.

Taylor, G.I. 1921. Diffusion by continuous movements. Proc. Lond. Math. Soc. 2:196-212.

Vanmarcke, E. 1983. Random fields: analysis and synthesis. MIT Press, Cambridge, Mass.

Van Genucthen, M.Th. 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci. Soc. Am. J. 44:892-898.

Van Genuchten, M.Th., and J.C. Parker. 1984. Boundary conditions for displacement experiments through short laboratory soil columns. Soil Sci. Soc. Am. J. 48:703-708.

Warrick, A.W., G.J. Mullen, and D.R. Nielsen. 1977. Scaling field measured soil hydraulic properties using a similar media concept. Water Resour. Res. 13:355-362.

Yeh, T.-C., L.W. Gelhar, and A.L. Gutjahr. 1985. Stochastic analysis of unsaturated flow in heterogeneous soils. I. Statistically isotropic media. Water Resour. Res. 21:447-456.

DISCUSSION OF PAPERS PRESENTED IN TECHNICAL SESSION 6, PART 1: SPATIAL VARIABILITY AND MODELING SCALE

David Bowles[1], Presiding
Chris Duffy[2], Recorder

PAPERS DISCUSSED

Effects of Spatial and Temporal Variability on Water Quality Model Development by D.A. Woolhiser, W.A. Jury, and D.R. Nielsen

Effect of Spatial Variability Upon Subsurface Transport of Solutes from Nonpoint Sources by G. Dagan, D. Russo and E. Bresler

SPECIFIC QUESTIONS AND COMMENTS

Question: (I. Moore, Department of Agricultural Engineering, University of Minnesota) Dr. Dagan, the underlying assumption of your analysis was that spatial variation is a random process. In some watersheds or some catchments, the spatial variation, say saturated hydraulic conductivity, is a random function, but in many other watersheds it is a function of topographic position which is a function of the evolution of the soils in the landscape. I was wondering if you could address that question.

Response: (D. Woolhiser, USDA-ARS, Tucson, Arizona) Yes, this is certainly the case that there are regularities in the soil physical properties, related to the position in the landscape, and this is very important in terms of the various mechanisms of say generating surface runoff, for example. I guess there is no question about this. We often have very sparse measurements, and so we have the error due to limited sampling of the measurement. We really don't know what happens between the measurement points and so I believe this is where the notion of the random field comes in. The variability that could occur between measured points, in fact, will certainly have an effect on flow or solute concentration found at the stream, or well, or whatever the case may be.

Question: (T. Richards, Cornell University, Ithaca, New York) Both authors, and several people during the past few days, have talked about some of the difficulties of taking measurements in the field. In particular, how do we deal with spatial variability, and what is the possibility of getting useful, integrated measurements on a scale which is of importance? I wonder if both of you might mention the possibility of using space averaging and looking at inputs from systems, to try to get a better sense of what the ultimate effects of variability are. Are we at a point yet where we can start to understand some of that variability from those averaged or integrated outputs?

Response: (G. Dagan, Tel Aviv University, Tel Aviv, Israel) I tried to emphasize this point in my talk. Perhaps I can be more concrete. Let's say your purpose is to find out what happens in an isolated observation well. This would be a kind of point value prediction. And at the other extreme let's say we have an aquifer and its outflow to a stream and we want to know the average, or the mass of solute which is transported into the stream. We don't really want to know what

[1]David Bowles, Professor and Associate Director, Water Research Lab, Utah State University, Logan, Utah.
[2]Chris Duffy, Department of Agronomy, Cornell University, Ithaca, New York.

happens at each point of contact, but rather over the entire extent of the aquifer. Now in terms of modeling and prediction, these are two different problems. From my perspective, the main difference is that in the first case the outcome of the model is subjected to a large degree of uncertainty, since at an isolated well all we can say is that the concentration will have an expected value and some standard deviation. Generally this uncertainty will be quite large, which means that our prediction may be of limited usefulness. In contrast, spatial variability of the output of a large scale aquifer has a smoothing effect, and the outflow is predictive. Although the expected value and thus the uncertainty is not so large, we still have parametric uncertainty. Even if you have a homogeneous aquifer, forget about spatial variability, you can predict the outcome. By the way, this is in line with what Chris Duffy has presented. He talked about such an integrated value, and he found if he modeled it with either spatially variable or just homogeneous information, you get more or less the same results. But, the results definitely depend on the homogeneous information, the transmissivity, or hydraulic conductivity. It's one number that we'll never know with certainty. So, we have this parametric uncertainty, but the result is still highly predictable. So to summarize, space average has a smoothing effect, and has the effect of reducing the impact of variability and uncertainty.

Response: (D. Woolhiser) In terms of the spatial averaging, we may consider a couple of factors. First, in terms of spatial averaging of inputs to a watershed, say, for example, precipitation. Spatial averaging is necessary as we go to larger and larger watersheds. Rather than taking into account all the spatial variability of rainfall, we do some averaging and determine an effective rainfall over the whole basin. However, if we use this average information to try to estimate the nonlinear infiltration process, severe biases result. The same would be true if we average infiltration parameters. If we made several measurements, for example, with a ring infiltrometer, we know that, from the field measurements made here at Utah State, there is tremendous variability and apparently the correlation scale is very short. If we averaged these and tried to use the results in a predictive sense, indications are that unless we are working on the really large events that we are going to have some severe biases in our predictions.

Another type of averaging, of course, is temporal averaging, and this is often done in some of the analytic studies looking at interaction of rainfall with infiltration. And again, because of the nonlinearities, this leads to some very severe biases. If we average the intensity process of rainfall, we'll always underestimate the volumes, we'll always underestimate the peak rates. Gideon referred to this question of equivalence as well. I think this is very important. If we can find simpler systems or simpler formulations that are somehow equivalent to the more detailed ones by some objective method, and we find it doesn't make any difference, then this is a very useful approach. We realize that this may work with some things but not with others. For example, we are looking at some equivalences in just surface runoff and we find that we can represent the watershed processes of a very small scale and then somehow aggregate these up to larger and larger scales, and get quite good estimates of our surface runoff response. But, the processes of erosion in the chemical transport may not still be equivalent so that is a difficult question.

Response: (G. Dagan) I would like to say one more word. If we accept the idea that when we average over a very large scale, we are smoothing and we reduce uncertainty, then the spatial variability is reflected by effective properties or an integrated equivalence. If we are able to determine those effective properties, then we can predict and this makes things much simpler because we replace the complex system by an equivalent or fictitious homogeneous one.

Comment: (G. Birch, Bureau of Rural Resources, Australia) Just picking up what Ian Moore said a little bit earlier, I think the point wasn't properly appreciated. Ian and I have done some studies on one catchment which is about 80 hectares in size. In another area I have done studies on smaller catchments and then used the results from that catchment to go up to a basin of about 40,000 hectares in size. In all our studies you have a finite depth of hydrologically active soil. On the 80 hectare catchment, for instance, we measured on a grid pattern the hydraulic conductivity

to estimate transmissivity for the catchment. The geometric mean was found to be very close to the analysis that Emmett O'Loughlin has developed for the development of saturation zones and runoff in this catchment. It was remarkable just how close the distribution of measured values was to the estimated discharge behavior of that catchment over a wide range of storm events.

Question: (D. Woolhiser) Yes, I think it was very interesting work. I have one question of the questioner if he would remain there. The generation mechanism from this then was primarily saturation, overland flow, or was it flow through the soil layers to the gullies and so on?

Response: (G. Birch) The analysis really is one which allows you to predict where saturation will occur in the landscape. It can occur, of course, in many positions but especially around drainage lines in the base of the catchment, and expanding up hillslopes where you have concave surfaces. Indeed the appearance of saturation is the primary mechanism by which discharge is generated. The importance of the analysis is that if you can correctly represent where saturation is occurring, and the rate of expansion of those saturation zones, then indeed you get a very good prediction, at least in our instances, of the discharge behavior both in peak and total amount. by modeling areas that were forested and areas that were cleared, I could represent the hydrograph of the much larger basin using the same principles, and with no further parameterization, from the smaller catchments.

Question: (T. Prato, University of Idaho, Moscow, Idaho) I think I'm struggling a little bit here because I'm wanting to hear something said about how we decide the scale that we should be working with here. I'm thinking about my own experience with geographic information systems where we're looking at erosion as well as water quality. For GIS you are looking primarily at the surface. The way I have looked at spatial variability is, for example, to look at the Universal Soil Loss Equation which tells you that the slope has a very significant impact on erosion. If you are in an area that has undulating topography, you want to pick your scale to be fine enough so that you can pick up differences in slope that are going to affect the erosion rate. And now I'm thinking that why not talk about three-dimensional geographic information systems where you get information below the surface. Could you comment on the potential uses of geographic information systems, and what is the appropriate scale of observation?

Comment: (D. Bowles) Maybe I can sort of piggyback onto your question; the issue of the role of remote sensing as a source of information.

Response: (D. Woolhiser) I expect there are other people in the room that could give you a much better answer than I can on this. I know there is a great deal of interest in applying GIS systems, geographic information systems, to develop input information for hydrologic models. Certainly from the standpoint of land use, the topographic information includes slopes at various scales. This is necessary for some of our complex models. One of the big problems is preparing the input for that type of parameter. In terms of getting below the surface, I am simply going to have to plead ignorance. Someone else may be able to address that question. The question of appropriate scale is one that we are struggling with. We are trying to look at it at Tucson with respect to surface runoff. A graduate student of mine is working on this and having some success in modeling at very small scales, and then objectively aggregating subareas while trying to preserve certain characteristics. But when you put erosion on top of that, or chemical transport, all bets are off. They seem to be much more difficult. We have known for a long time that you may do a good job of predicting a hydrograph if it doesn't really matter how that water gets to the stream. We may say it is surface runoff or it is groundwater, but when we deal with the chemical transport and erosion, it matters a great deal how it got there.

Question: (D. Gustafson, Monsanto, St. Louis, Missouri) I would like to ask Professor Dagan: You mention that spatial averaging will allow you to ignore some of the spatial variability, and your arguments were based on length scales where your outflow is much greater than the length

scale of your input source. Isn't it true that one of the primary uses of these types of models is to predict concentrations in wells? After all, what we are really trying to do, in many cases, is estimate human exposure. Certainly in the mid-west United States we have thousands, if not millions, of these wells from which people are withdrawing water and being exposed to agricultural chemicals. And so, a very important issue is trying to predict the distribution of likely concentrations in a series of points rather than some spatially averaged number which may not characterize the human exposure at all. I wonder if we are painting a little bit too rosy of a picture in terms of how much this spatial averaging helps us in what we're really trying to use these models for.

Response: (G. Dagan) If you have a body of wells which are discharging in a kind of communal system, I guess that what you are after is not really the output of the individual well, but again a kind of average of a few bodies of wells. Again, in terms of model prediction we can handle this with smaller uncertainties. For point sources, this problem of prediction at a certain point, the concentration is averaged over a very small volume. And what we come out with is large uncertainties. My point is that if you are ready to accept a result which concerns a body of wells or a larger area, then the model will tell you something more precise, less affected by uncertainty. In principle the model can do either job. The management question is, what do you do with the numbers?

Question: (G. Vachaud, University of Grenoble, France) I'm shifting from wells to soil, and I would like to raise a question to Professor Dagan concerning uncertainty of solute concentration measurements in soil. When we consider that a set, like the one you have used as an illustration, and when we analyze this set, do we obtain results which are related to the intrinsic state of the soil concentration, or do we add information which is in fact related to variability of the measurement technique? And by that I wonder if we are interested in the spatial analysis to obtain the spatial property of the experimental technique, which is, of course, associated with experimental bias. In other words, what are we measuring when we take a solution from the soil, and do we get something which is more or less related to a real concentration?

Response: (G. Dagan) Well, I think you can answer the question better than I can because you are dealing with measurement. But what I can tell you from the modeling perspective is, we have to attach to measurements as well, a certain level of uncertainty. And if we can quantify the relationship between measurement and the real quantity, then we can use that input. In the picture I presented, the assumption was, that what was measured is what is really there. But the methodology exists to incorporate measurement errors into the modeling effort, but then it becomes a circular kind of argument. How do you know what is a measurement error if you don't really know what happens.

Comment: (D. Woolhiser) Just one comment. I'm certainly not an expert in measurement, but I did run across a very interesting paper in my review by some scientists from Israel who used a micro technique of measuring concentrations above or near a water table of a polluted aquifer in Israel. They found very high variability in concentrations at the scale of centimeters. If that is the case, we certainly have some great variability.
Question: (Audience) Some of the concentration data from recent field experiments by Frieberg and others is shown to be very smooth and regular, and you could almost approximate the soil as if it were almost uniform in terms of hydraulic conductivity. When you compare that to the profile you showed of heterogeneity in the vertical direction of hydraulic conductivity, it doesn't quite make sense. Can you comment on the concept of chemical observations being fairly predictable using a gross approach to hydraulic conductivity, compared to looking at very detailed hydraulic conductivity variations?

Response: (G. Dagan) I have seen these maps of measured values, and they don't look so uniform to me. There are new papers that are a kind of data bank for anybody interested in analyzing

them, where there are thousands of measurements. I think in the paper by Barry and Sposito you will see they are quite variable in space. Not only in the horizontal plane but also in the vertical direction. Frieberg and others have analyzed the experiment using theoretical approaches with good results. But what they did is a very interesting point you have to keep in mind. They took the concentration to be a kind of solid body, that is already averaged over the vertical. But you can see the distribution itself is quite irregular. There are spots of high concentrations. They calculated the moment of inertia of this body, and assumed that it can be replaced by this irregular distribution which looks something like an ellipse. This is what you get from a Gaussian model. If you start with a pulse, you get an ellipse. They compared the moment of inertia, the spread of that ellipse, with the spread of this cloud, and they found a good correspondence. But the correspondence was not for concentration at each point here. This cannot be predicted. The prediction was for a gross feature of this solid body. So this brings us back to the same problem, the scale of averaging. We don't want to know the concentration at each point in such a complex pattern, but we can predict the spatial moment of this body which gives you a gross picture of its spread. And in that we can do a much better job than in predicting the point value.

COMPLEXITY AND UNCERTAINTY IN PREDICTIVE MODELS

K.J. Beven[1] and A.J. Jakeman[2]

ABSTRACT

This paper addresses the problems of predicting the responses of complex real-world systems about which we have only limited information. Sensitivity analysis and the use of stochastic models are considered as ways of reducing parameter dimensionality. The importance of estimating predictive uncertainty is illustrated through two studies of surface water contamination.

"To prophesy is extremely difficult, especially with respect to the future" (Chinese proverb).

ON THE NATURE OF PREDICTION

Much of the current research work in water quality problems is concerned with developing predictive models. The aim of this activity is generally to provide the capability of assessing the impact of natural or man-induced perturbations to a system of interest on the quality of the water resource. We will reserve the term prediction for situations where it is not possible to make measurements. Most models, however, do not get used for prediction. Ignoring that vast number of models that are only ever used to make predictions about the behavior of hypothetical systems, those that do address real systems are most frequently used to imitate the past behavior of a system during a period for which observations of system behavior are available, an activity which we will refer to as "postdiction" (and has also been called "hindcasting"). Postdiction is a valuable way of showing that a model can indeed reproduce, with some more or less acceptable degree of error, the behavior of a system of interest. It is a necessary part of model development, required to demonstrate that a model may be suitable for use in prediction.

Prediction, however, requires extrapolation in time, condition, and/or space. We may need to predict behavior in the future or under extreme high or low flow conditions, or unmeasured system states at different internal points, or the behavior of similar ungaged systems at different locations or in different physioclimatic regions. These are all situations where the degree of error in reproducing system behavior is essentially unmeasurable. We have then no guarantee that the magnitude of the modelling errors that we obtain during postdiction will be similar when the model is used for prediction. It should therefore be a requirement of water quality scientists that in making predictions, they should also attempt to estimate the uncertainty associated with those predictions. This paper is concerned with the relationship between model complexity and uncertainty in our attempts to predict the behavior of the complex natural systems involved in water quality problems.

THE SEDUCTIVENESS OF COMPLEXITY

In constructing a predictive model of a water quality system, it appears attractive to the scientist to build in as much physical and chemical understanding as possible. The progress of understanding of water quality problems over many years has resulted in a large body of knowledge about the

[1]K. J. Beven, Reader in Hydrology, Centre for Research on Environmental Systems, Institute of Environmental and Biological Sciences, University of Lancaster, Lancaster, LA1 4YQ, UK.

[2]A. J. Jakeman, Senior Research Fellow, Centre for Resource and Environmental Studies, Australian National University, Canberra, ACT 2601, Australia.

operation of such systems that can be expressed in a mathematical form. We can readily conceptualize sets of descriptive equations of processes that involve several space dimensions and numerous state variables and parameters that may vary in space and time. We know that the model equations may be validated and parameter values varied by careful experiment. It is a common assumption that such "physically-based" models must be inherently more accurate than simpler models.

Such an assumption is often not justified. The problem is twofold. There is evidence that even some of the best documented physically-based descriptive models may not properly describe the nature of the processes operating in field situations. An example relating to dispersion in rivers is described below (see also discussions in Beven 1987, 1989). In addition, a combination of budget limitations and the uncertainties and variability in boundary conditions and other uncontrolled variables means that there may be no possibility of measuring all the data required to calibrate and properly validate a complex "physically-based" model.

The modeler then faces a dilemma. On the one hand, there is a desire to ensure that a predictive model incorporates, as far as possible, our physical understanding of the processes involved; on the other, the complexity of such a model may result in a structure that is vastly overparameterized with respect to the possibilities for calibrating its parameters with the data available. The attraction of a fully physically-based model is seductive, in that the physical knowledge on which it is based should allow a greater faith in its predictions, particularly when predictions are required outside the range for which a model has been calibrated. We know however that if a model is overparameterized then calibration is bound to lead to a high degree of uncertainty in the values of those parameters and consequently to greater uncertainty in the predictions (see also Beck 1983).

How then to resolve this dilemma? It is quite clear from the literature that different modelers adopt different strategies. At one extreme, some make deterministic "physically-based" predictions in blind faith that their simulation of physical processes will lead to accurate predictions even if some parameters have to be fixed or estimated a priori and others calibrated with only few observations. At the opposite end of the spectrum, others use "black-box" models whose level of complexity is justified entirely on the measured data available and make little reference to physical processes. We have learned from these latter techniques that the responses of complex natural systems can often be represented by models of low order and few parameters that embody the dominant modes of behavior. Such models can be robust in terms of the uncertainty associated with calibrated parameter values and predictions. They have the advantage that the linear systems methodologies usually used have well developed techniques available for estimating predictive uncertainty. They have the disadvantage that in some cases the extrapolation of parameter values and associated uncertainty estimates in time or space for predictive use may be impossible.

The choice of an appropriate predictive technique may then require a compromise between building in sufficient physical knowledge and avoiding an overparameterized model. This involves a reduction in dimensionality of the modelling problem, perhaps in terms of the number of space dimensions from a fully three-dimensional representation to a one-dimensional or even fully lumped structure, but most importantly in terms of the parameter space, the number of parameters to be calibrated or estimated. Several techniques are available for reducing dimensionality and these will be considered below.

REDUCING DIMENSIONALITY: SENSITIVITY ANALYSIS

There are many forms of sensitivity analysis and definitions of sensitivity coefficients (see for example McCuen 1976, Frank 1978, Rabitz et al. 1983, Hornberger and Spear 1981). For a predictive model that is defined by a general system of differential equations of the form

$$\frac{\delta c_i}{\delta t} = F(\underline{v}, \underline{\alpha}, \underline{x}) \qquad [1]$$

where c_i = the i th state variable of interest,
$\underline{\alpha}$ = the vector of parameters,
$\underline{x}$ = the space dimensions, and
t = time.

Then elementary sensitivity coefficients may be calculated as the gradients

$$\frac{\delta c_i}{\delta \alpha_j}; \quad \frac{\delta^2 c_i}{\delta \alpha_j \delta \alpha_k}.$$

The sensitivity coefficient is a reflection of the dependence of a state variable change in a model parameter. Evaluation of these gradients may be achieved analytically for some simple models (see for example Beven 1979) but for most cases of interest it will be necessary to use numerical approximations. Parameters to which model predictions are sensitive will show large values of the gradient terms; those having little effect will show small values. Sensitivity information can be used to provide valuable insight into the operation of a predictive model and, by analogy, into the dominant physical and chemical controls of a water quality system. It can be a valuable aid to reduction of parameter dimensionality in that it suggests which parameters can be eliminated from the calibration process.

A number of problems arise in this form of sensitivity analysis. A primary problem is that the sensitivity coefficients have to be calculated at a particular point in the parameter space. Normally, this is at some "optimal" set of parameter values determined as a result of a calibration exercise. In models containing a large number of parameters, however, there may be many sets of parameter values that would give nearly equivalent goodness of fit to the observed data (the equifinality problem), although all will be in error to some greater or lesser extent. The ranking of sensitivity coefficients for different parameters may well vary when they are evaluated for different sets of calibrated parameters. In addition, in a time-varying simulation the sensitivity coefficients (and their rankings) can be expected to vary over time in response to changing boundary conditions and system dynamics. Finally, both the boundary conditions input to the model, and the measurements on which calibration of parameters is based will both be subject to error, which may affect the estimated sensitivity coefficients.

An alternative approach to evaluating sensitivity that overcomes some of these difficulties has been suggested by Hornberger and Spear (1981). Their "Regionalized Sensitivity Analysis" (RSA) has been used in the context of water quality modelling by Spear and Hornberger (1980), Whitehead and Hornberger (1984) and Hornberger et al. (1985, 1986). The RSA procedure makes no attempt to identify any optimal set of parameter values, and allows that parameter values may be subject to considerable uncertainty. It also makes no assumptions about the shape of the parameter response surface. The method is based on Monte Carlo runs of the predictive model of interest using sets of parameter values chosen randomly from distributions specified a priori. These distributions are chosen to reflect the degree of uncertainty with which any particular parameter can be estimated. Any known correlation or functional dependence between parameter values can be preserved. The procedure depends crucially on a quantitative or qualitative definition of acceptability of a particular run as a simulator of the system of interest. On the basis of this definition each run is assigned to either "behavior" or "non-behavior" classes. The distributions of parameter values contributing to each of these classes are then compared statistically. If they are significantly different for a certain parameter then it may be concluded that the simulations are sensitive to the value of that parameter.

Regionalized sensitivity analysis has been used by Hornberger et al. (1985) as a means of reducing the number of parameters in a hydrologic model. They show that the resulting model structure may depend, however, on the criterion (in their case different forms of objective function) used to define behavior and non-behavior. They also suggest that there may be a rather subtle interdependence between "sensitive" and "non-sensitive" parameters and that it may be better to eliminate model parameters by reducing the complexity of the model structure rather that fixing parameter values to be constants.

REDUCING DIMENSIONALITY: THE ADVANTAGES OF STOCHASTIC MODELS

When processes are complex, it may be difficult to find a satisfactory deterministic model. When data are available an alternative strategy is to use a stochastic approach which assumes the attribute of interest at time k is a sum of two components: a function of the observation history and process parameters until time k-1 and a purely random component which may be acted upon by another function. As we shall see, stochastic models need not be purely black-box in terms of their level of process understanding. Indeed, any deterministic model can be reformulated as a stochastic model. For example, partial differential equations can be written in state-space form with the state evolution component representing the deterministic model and the observation equation allowing for stochastic errors if desired.

The advantage of stochastic models is that they can be assessed according to formal criteria for their statistical applicability to available data sets. And, inferred parameter values, lag and error structures and associated model properties may also be tested for physical realism. Alternatively, prior knowledge including model structure and the distribution of parameter values can also be injected into stochastic formulations (e.g. Schweppe 1973).

In the remainder of this section we consider a particular class of stochastic models known as linear difference equation or transfer function models. They are used here partly to demonstrate some of the properties and conditions of the application of stochastic models. The main reason for their introduction here is to argue that linear difference equation models cover a wide variety of model types, and depending on the objectives of the modelling exercise and the system under study can be useful in their own right either as final model forms or as exploratory mechanisms for further understanding of the dominant system relationships.

A class of linear difference equation models (Box and Jenkins 1970, Young 1984) can be defined at any time step k as

$$x_k + a_1 x_{k-1} + \ldots + a_n x_{k-n} = b_o u_k + b_1 u_{k-1} + \ldots + b_m u_{k-m} \qquad [2a]$$

$$\epsilon_k + c_1 \epsilon_{k-1} + \ldots + c_q \epsilon_{k-q} = e_k d_1 e_{k-1} + \ldots + d_p e_{k-p} \qquad [2b]$$

$$y_k = x_k + \epsilon_k \qquad [2c]$$

The variables in this formulation are the input u_k and noise-free output x_k in the deterministic component equation 2a, the model residuals ϵ_k and the white innovations of the noise model e_k. The parameters to be estimated are the system model parameters $\underline{a} = (a_1, a_2,\ldots,a_n, b_0, b_1,\ldots,b_m)^T$ and the noise model parameters $\underline{c} = (c_1, c_2,\ldots,c_q, d_1, d_2,\ldots,d_p)^T$. Of course, the model can be written in multivariate forms (Young 1984, Jakeman and Young 1979, Jakeman 1979, Jakeman, et al. 1980) and that spatial discretization of a variable can be accommodated by vector notation (Bennett and Chorley 1978). For convenience, however, we will continue to use the scalar notation introduced in equation 2.

Equation 2 covers a number of model types including stochastic linear state-space models (see Young 1979, and Jakeman and Young 1979) and the autoregressive - moving average (ARMA) models considered by Box and Jenkins (1970). The latter is the purely stochastic case where the deterministic component of equation 2 describing the evolution of x_k is omitted. The class of ARMA models, where behavior is characterized according to past statistical patterns in the attribute of predictive interest, has proven generally useful for forecasting, especially when the forecast is over small lead times. The collection of papers in Anderson and Perryman (1982) provide an indication of the interest in and wide variety of problems tackled by this approach.

An added flexibility of largely stochastic representations such as equation 2 is that strong linearity of all the true underlying relationships need not be assumed. Complexity and nonlinearity can often be reduced by aggregation of model structure, at the same time leaving sufficient explanatory detail in the model for it to be sensitive to the desired input variables. Young (1978) advocates a law of large systems which states that there are a small number of dominant behavioral modes which characterize many systems problems. Indeed, experience indicates that this law is often approximately linear. The applications in Jakeman (1985), Mahendrarajah et al. (1982), Jakeman, et al. (1980), Young, et al. (1980) and Young (1978, 1982) bear some testimony to this hypothesis. These examples encompass mainly surface and groundwater flow and solute transport problems and some air quality pollutant transport problems.

Transformations of variables can often be utilized to change a nonlinear systems representation to a linear one upon which the analysis and modeling can proceed. For example, logarithmic transformations of time series data are frequently used in econometric applications. Long term low frequency trends may be removed from data and modelled separately.

Input variables may be defined as nonlinear functions of other variations to allow the basic equation to be linear-in-the-parameters. Rainfall-runoff models have been developed of the form equation 2 where the runoff y_k is determined from rainfall after modification for nonlinear evaporation and soil moisture effects to yield an effective rainfall term u_k (Whitehead, et al. 1979, Mahendrarajah et al. 1982).

Depending upon the objectives of the modeling exercise, it may be possible to construct models on subsets of the data where in each case linearity is a reasonable approximation. Jakeman, et al (1984) used models of the form equation 2 to predict missing stream flows which are output from a complex karst drainage system. Separate linear models are developed which are statistically valid for fixed portions of the data. Missing flows within such consistently-behaving hydrologic periods can be predicted from the statistical model which is identified from other data within the period. Of course, if the objective had required more understanding of the hydrogeologic system, then this approach would not have been sufficient.

Finally, algorithms for linear models of the form equation 2 can be derived for the case where one or more of the parameters is allowed to be time-varying (e.g. Young 1984). The additional complexity of the algorithms over the time-invariant case is minimal especially if recursive algorithms are applied. Recursive, as opposed to the usual and equivalent en-bloc, estimation adds a useful dimension to time series analysis. It allows updating of parameter estimates $\underline{a}_k$ and their associated covariance matrix P_k at each time step k as additional measurements become available. Thus $\underline{a}_k$ (and P_k) can be expressed as its value at the previous time step $\underline{a}_{k-1}$ plus a correction term based on the new input and output observations at time k weighted according to confidence in the model at time k.

In addition to allowing on-line real time forecasts, the major advantage of the recursive facility is that parameter estimates can be allowed to vary with time to best explain output behavior. The pattern of variation can then be analyzed to check linearity of the model or to identify and quantify inadequacy over any particular period of the data set. Hypotheses can then be inferred to

improve model structure and these can be tested with the aid of recursive algorithms and other diagnostic checks (see below).

As with any model, the resultant models identified from systems like equation 2 and its extensions are only strictly valid for the data sets upon which they are calibrated and validated. This validity is more conditional the more the model structure and parameters have a statistical or black-box structure as opposed to a recognized process interpretation. At one extreme, we can only expect ARMA models to predict well on new data sets if the new behavior retains an ergodic and stationarity property. At the other extreme, we might expect an accepted process model written in the form of equation 2 to predict well on future data sets if the model was well identified; for example the statistical parameters which have physical meaning were estimated with low variance, the output behavior was well explained with uncorrelated model residuals and so on.

SOURCES OF UNCERTAINTY

In order to reduce or just be aware of predictive uncertainty, it is helpful to consider the major sources. They can be placed within the following categories: properties of the data available; model structure identification; parameter estimation method; algorithmic implementation; verification and validation procedure; future inputs. Jakeman (1988) also discusses these sources of uncertainty and provides some illustrative examples.

Properties of the Data

Some of the sources of error are well appreciated. Sampling and measurement error affect the accuracy of any selected model as does the quantity and quality of data available for calibration. Just as important is the persistence of the excitation (see Young, 1984 for a mathematical definition) that is input to the system by the forcing variables. A system with two frequency responses cannot be identified from input-output data sets where the input data is of a single frequency. In more physical terms, the representativeness of the observation or measurement period with respect to the range of future conditions intended for model use is a factor in the certainty of model predictions. Finally, certain model choices and parameter estimation algorithms place restrictions on the type of data used; for example, data with stationarity or insignificant autocorrelation. Failure to meet such conditions affects predictive uncertainty.

Model Structure Identification

It is in this category that the most troublesome sources of uncertainty lie. Lack of necessary prior knowledge about basic processes including the values of physical variables can be a problem with natural systems. More often, however, the larger problem is the complexity of the system of interest, which in turn can be exacerbated by inadequate data to explore the determining relationships. Reduction of the system into component subsystems which are tractable in their definition will only be successful if some of the linkages or connections between subsystems can be defined and the remainder ignored. There is a trade-off to be achieved in the choice of model structure. There must be sufficient determinism in the model to explain the output behavior under future conditions but not too much as to lead to strong interdependence or ambiguity among model variables and parameters. Consequently, the level of aggregation of model structure and of variables in the space and time domains is important. Some level of stochasticity in the model will often be preferable either to characterize the aggregate behavior or the errors or both.

Turn now to the situation where the basic processes associated with the modelling problem are well understood under reasonable assumptions. While these assumptions should be appreciated as a source of uncertainty if they cannot be fully validated, there is no further discussion of that here. A less obvious but crucial consideration is the mathematical nature of the classical formulations

associated with these processes. They should not unduly present a source of uncertainty. A good example is the case of an ill-posed formulation where either a unique solution cannot be found or the solution does not depend continuously on the input data. The practical consequence of this property is that parameter estimates are highly sensitive to small changes, such as errors, in the observational data.

Ill-posed formulations are pervasive in hydrology where one attempts to identify physical properties of surface and sub-surface systems from indirect observations such as flow and solute concentration. These identifiability problems require that an ill-posed formulation be replaced with one which is well-posed by restricting the solution set. For example we may lower the order of a parameterization, impose smoothness constraints on it or propose a probability distribution for some of the parameters. If these restrictions are assumed rather than known, then it is useful to attempt to quantify the related uncertainty of the solution. Investigation of the effects of different prior assumptions designed to improve the identifiability is also worthwhile in this case. Dietrich, Jakeman and Ghassemi (1988) show the pervasiveness of ill-posed formulations in flow and solute transport applications and discuss strategies for their treatment.

Parameter Estimation Method

Model development most frequently involves some parameter estimation technique. Estimation should involve some choice of criteria such as minimization of an objective function to yield parameters with certain desirable properties. The level of uncertainty may well be substantially affected by this choice since different estimators can have different properties. Depending on the data and the objective, the choice may require some of the following properties: robustness (to outliers), unbiasedness, consistency, minimax, asymptotic efficiency, etc. (see for example, Young, 1984).

Algorithmic Implementation

A subsequent technical consideration for parameter estimation or for using a model in predictive mode once it has been developed is the choice of numerical algorithm. Sources of uncertainty here result from the limitations of most computers in that they do not work with all the real numbers but a finite subset, usually a floating point number system. Anderssen and de Hoog (1983) provide a lucid exposition of some of the problems in ensuring acceptable computational performance. Problems such as rounding and truncation errors, instability and poor convergence are not as pervasive as a decade or so ago with many 'safe' routines now available in packaged form. However, there is still a great deal of user flexibility in the choice of algorithm which may affect uncertainty. At the simplest level, one may have to choose the value of some input for a packaged subroutine, for example, the implicitness of a finite difference scheme. A more pervasive uncertainty which occurs frequently in natural system problems, whether or not packages are available, is that sample sizes may be sufficiently small that numerical performance may not be known and will need assessment through construction of synthetic problems with known solutions (see for example, Jakeman and Young 1983).

Verification and Validation Procedure

If diagnostic checking of assumptions is not sufficiently exhaustive then this may lead to over-confidence in uncertainty measures. The level of information and data available for validation also affects the certainty of a model since, in a strict sense, a model is only valid for the conditions represented by the calibration and validation data sets.

Future Inputs

To use a model in predictive mode may require some future input to the model such as a time series for a forcing variable, a probability distribution for a variable or a parameter, or some specific parameter value. If we are merely interested in the sensitivity of the model predictions to different input scenarios chosen arbitrarily, or the input is controllable, then no uncertainty enters here. On the other hand, if the prediction objective requires knowledge of the causal inputs driving the model, then the contribution to uncertainty from this category can obviously be enormous. The message is to try to avoid developing a model structure which requires knowledge of an input that is too difficult in itself to predict.

ASSESSING PREDICTIVE UNCERTAINTY

All these sources of uncertainty will lean to uncertainty in our estimates of the values of model parameters and in the resulting model predictions. It has been argued above that it is important to attempt to assess the magnitude of that uncertainty in a realistic way. In the case of linear systems an estimation technique can usually be found which yields the probability distribution of parameter estimates, asymptotically and under suitable regularity conditions. If we consider the linear transfer function or difference equation model 2 then we have the result below. For other linear model representations, similar results and conditions apply.

Theorem (Jakeman and Young 1983, Pierce 1972)

If in equation 2

(A1) the e_k are independent and identically distributed with zero mean and constant variance;
(A2) the parameter values are admissable; and
(A3) the input u_k is persistently exciting;

then the maximum likelihood estimates of the parameters $\underline{a}$ and $\underline{c}$ possess a limiting normal distribution.

In other words, associated with an estimation algorithm is a central limit theorem arguing asymptotic normality of the parameter estimates. In practice, to obtain the parameter covariance matrix and hence their distribution, it is then assumed that

(A4) strictly unknown quantities in the expressions for the covariances, such as the true parameter values and noise-free system output, can be approximated by the estimated values.

Another assumption relates to the question of how large is asymptotic for the result to hold, that is

(A5) the sample size used for estimation is suitably large for convergence to the first and second moments of the asymptotic normal distribution.

Some comment on the applicability of these assumptions in application is warranted. The conditions (A1)-(A3) are not unduly harsh. If these conditions are not satisfied, then poor estimation results will be obvious in any case. Further, they are easily verified. For linear systems, (A4) is not limiting. At least small changes from the true parameter values do not yield large changes in estimates of the covariances. Extensive simulation work is required to prove this for single-input single-output systems, see Young and Jakeman (1979) and for multivariable systems

see Jakeman (1979), Jakeman and Young (1979) and Jakeman, et al. (1980). Fortunately, assumption (A5) is also not too limiting. The latter references also show that sample sizes of the order of 100 are adequate. Indeed, for low system parameterizations of say 2 to 4 parameters, sample sizes of around 50 can be adequate and this is certainly the case for inputs with a high level of persistent excitation.

However, in practice the major limitation is:

(A6) the underlying system can be represented to a good approximation by the linear model for all uses of the model.

This can and should be partially investigated with diagnostic checks such as:

(D1) the model residuals ϵ_k are not cross-correlated with the input $u_{k-\tau}$ at lags $\tau = 0, 1, 2, ...$;
(D2) the innovations e_k of the error model (2b) are not auto-correlated; and
(D3) the parameter estimates are time-invariant.

These are necessary but not sufficient conditions for the relevance of the linear model. Unfortunately a linear model applied to a given data set may explain the output behavior to the satisfaction of criteria like (D1) - (D3) yet a situation can easily be envisaged where model errors are hidden. For example, omission of an important factor in the input may not be obvious because of the ability of the statistical estimation procedure to calibrate the deficiency by sympathetic adjustment of a parameter value. It is only by physical insight and testing on other data sets where the inadequacy may be more obvious that such models can be improved. Because of the assumptions (A1) to (A6), and especially (A6), the algorithmic estimate of the covariance matrix of parameters should always be regarded as a minimum estimate.

But what if we cannot invoke an estimation procedure which yields such nice properties? For example, any one of the following limitations for an asymptotically efficient algorithm may be a crucial impediment to a modeling exercise:

(L1) the algorithm is too computationally demanding;
(L2) it is unstable on small sample sizes;
(L3) the error assumptions need to be relaxed so that a more robust algorithm is required to handle outliers or model errors uncorrelated with the input.

In such cases, and also in the case of nonlinear models, other methods must be adopted to attempt to characterize predictive uncertainty. No general theory equivalent to the linear case exists and approximate methods must be used. A number of different classes of techniques have been used, including first order analysis (e.g. Garen and Burgess 1981); the Rosenblueth method (Rosenblueth 1975, Guymon et al. 1981, Rogers et al. 1985), Monte Carlo methods (e.g. Gardner et al. 1980) and bootstrapping (Efron and Tibshirani 1986).

Of these methods, the most general is the Monte Carlo approach, in which a number of model runs are made using random selections of uncertain parameter values or boundary conditions. The Monte Carlo technique is not limited by the degree of nonlinearity of the model, or the degree of uncertainty or assumptions about the form of the distributions from which the random selections are made. Any cross-correlation between the chosen random variables can be preserved. The major disadvantage of this approach is computational expense, especially for complex models, since a large number of runs may be necessary before convergence of the predictive uncertainty estimates is achieved. However, it is likely that this method will become more popular as the computational power of desktop workstations increases, because of the advantages of being highly flexible, easily understood and readily implemented.

APPLICATION OF A PHYSICALLY-BASED MODEL FOR SALINITY MANAGEMENT

In this section we outline results of a study involving the development of a model for salinity management. The aim is to make the following points:

(1) the type of model and in particular its level of physical detail should depend upon the knowledge available (including system and data) and the objectives;
(2) simple liner time-varying models employed with recursive estimation algorithms have a number of uses in modeling;
(3) linear-in-the-parameters systems models can be physically plausible and predict well on data sets other than those upon which they were calibrated;
(4) within a linear stochastic framework, identification of model structure, estimation of parameters and diagnostic checking become straightforward; and
(5) uncertainty measures can be approximated.

Background: Problem, System, Data and Objectives

Rising saline groundwater tables influenced by deep rainfall percolation after land clearing and exacerbated in some areas by subsequent irrigated agriculture is the dominant contributor to salt gain in Australia's largest river system, the Murray-Darling (see Thomas and Jakeman 1985). Effective management of salinity in the river calls for information about the interaction between the river and groundwater systems under a range of conditions: from upstream discharge and salinity concentrations; groundwater inflow and its salinity; stream volume gains such as return flows; and stream volume losses, such as channel losses, evaporation and diversions.

Basic information for characterizing this interaction can be obtained from distributed-parameter models which express the conservation of solute and water mass in the form of partial differential equations. However, in many situations the data required to calibrate such models with sufficient accuracy are not available. Accurate stream discharge and solute concentrations may not be available on a sufficiently intensive basis. The associated groundwater inflow will almost always need to be estimated indirectly and its solute concentration may or may not be known. The need to estimate stream volume changes will vary depending on the season (e.g. whether evaporation or irrigation diversions are significant) and the reach of interest.

In addition, the application of finite difference schemes to these distributed-parameter models, where advection dominates, may lead to numerical instabilities. Although finite element schemes may reduce numerical problems, it is at the cost of increased computational complexity (Taigbenu and Liggett 1986).

On the other hand simplifying assumptions often leading to lumped parameter systems, may allow the modeler to derive an approximate description of the system of interest which may be satisfactorily validated with the data available and remain useful for management purposes. Utility often required retaining the kinematic wave nature of advection so that prediction can be accurate on a time scale consistent with the sampling interval for the observations.

With this philosophy, Dietrich, Jakeman and Thomas (1988) consider the following problems:

(1) quantification of groundwater salt load and inflow from daily upstream and downstream concentration measurements,
(2) prediction of daily stream solute concentration from upstream discharge and solute concentration, and groundwater accessions to the stream.

Basic Equations

The starting point is conservation equations of water and solute mass. Neglecting diffusion and assuming all relevant stream properties are averaged over the wetted cross-sectional area A, these are (Thomann 1972)

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = q - E - F \qquad [3]$$

$$\frac{\partial Ac}{\partial t} + \frac{\partial Qc}{\partial x} = qs - Fc \qquad [4]$$

where
- A = the wetted cross-sectional area, $[L^2]$
- Q = the stream discharge, $[L^3T^{-1}]$
- q = the positive inflow per unit length from the aquifer into the river, $[L^2T^{-1}]$
- s = the aquifer solute concentration, $[ML^{-3}]$
- E = evaporation rate per unit length, $[L^2T^{-1}]$
- F = the net loss per unit length of stream volume (diversions and channel losses), $[L^2T^{-1}]$
- c = the concentration of stream solute, $[ML^{-3}]$
- x,t = the distance [L] and time [T], respectively.

Note that all variables may depend on both x and t. Inserting equation [3] in [4] yields

$$\frac{\partial c}{\partial t} + u\frac{\partial c}{\partial x} + \frac{q-E}{A}c = \frac{qs}{A} \qquad [5]$$

with $u =$ Q/A being the averaged cross-sectional advective velocity. $[LT^{-1}]$

Equation 5 is a hyperbolic first order partial differential equation easily solved by the method of characteristics. The solution, given by Dietrich, Jakeman and Thomas (1988) is

$$c(k) = \Omega(K)\left[\int_{k-\tau_k}^{k} \frac{qs}{A\Omega(t)}\,dt + c_0(k-\tau_k)\right] \qquad [6]$$

with

$$\Omega(t) = \exp\left[\int_{t-\tau_k}^{t} \frac{q-E}{A}\,d\theta\right] \text{ for } k-\tau_k \le t \le k$$

where
- τ_k = the travel time of a solute parcel arriving at the downstream reach end at time k,
- $c(k)$ = the time k downstream-end solute concentration $[ML^{-3}]$, and
- $c_0(k)$ = the time k upstream concentration, $[ML^{-3}]$.

Throughout this example the time k is in days. The integrals are curvilinear, being evaluated at point (x(t),t) along the characteristic curve Γ_k.

To apply equatioh 6, the quantities q, s, E and τ_k need to be known or assumptions need to be made. τ_k can be found as a function of discharge by one of two models given in Dietrich, Jakeman and Thomas (1986).

Model for Quantifying Salt Load and Inflow

Consider some simplifying assumptions which we will adopt for both our objectives: evaporation is negligible, groundwater solute concentration remains constant and is large compared to c_0, $A = fQ^g$ with g known through a stage discharge curve. Equation 6 can be approximated by

$$y_k = \alpha_k u_k + \epsilon_k \qquad [7]$$

where y_k = $c(k) - c_0(k-\tau_k)$ is the solute concentration $[ML^{-3}]$ added within the reach,
α_k = sq_k/f is an unknown parameter varying with inflow at time k, $[MT^{-2}]$; ϵ_k has been added to characterize all errors, $[ML^{-3}]$ and

$$u_k = \int_{k-\tau_k}^{k} \frac{dx}{Q}$$

If f is known from a stage discharge curve, then knowledge of α_k yields knowledge of salt load sq_k. An estimate of α_k can be obtained from data on y_k and u_k.

Data for y_k can be found from use of the simple routing and travel time models of Dietrich, Jakeman, and Thomas (1986). These predict c(k) from $c_0(k-\tau_k)$ when salt is conserved within the stream, that is with no groundwater inflow. Also data for u_k can be obtained if there is reasonable information on Q, but consider for simplicity the case of inflow within the reach being a point source. Then $u_k = 1/Q_k(x_0)$.

It is preferable to use smoothing algorithms to obtain α_k in equation 7 from y_k and u_k in order to eliminate the effect of errors, ϵ_k. Jakeman and Young (1984) and Young (1984) provide simple recursive filtering and smoothing algorithms using Gauss-Markov assumptions for α_k, viz.

$$\alpha_k = \phi \alpha_{k-1} + e_k \qquad [8]$$

with ϕ, a known transition matrix, and e_k, the Gaussian innovations. Thus, algorithms such as equations 7 and 8 can be used to quantify the salt load and indeed the inflow of s has been measured. We leave this simple case now and look to finding a more deterministic model explaining groundwater inflow and hence allowing prediction of downstream solute concentration with changing inflow q_k.

Model for Predicting Downstream Solute Concentration

To derive a sub-model for q_k we can refer to properties of analytical solutions to the equation of flow in an aquifer perfectly connected to a stream and to Darcy's law (Bear 1972). These solutions show that the water pressure and river stage height h are related by a linear Volterra convolution integral. Analytical solutions have been given for special cases, none of which apply to the aquifer under consideration here. An alternative adopted by Dietrich (1988) is to hypothesize a convolution of the form

$$q(t) = \int_0^t h(v)\, K\,(t-v)\, dv \qquad [9]$$

and attempt to identify the impulse response K indirectly from the available stream-aquifer data. Note that h must be specified with respect to reference coordinates at a distance from the aquifer where groundwater pressure is zero. The equation can be approximated for discrete data using a rational polynomial representation of K, that is

$$\frac{B(z^{-1})}{A(z^{-1})} = \frac{(b_0 + b_1 z^{-1} + \ldots + b_m z^{-m})}{(1 + a_1 z^{-1} + \ldots + a_n z^{-n})} \qquad [10]$$

Model equation 7 can then be written

$$y_k = \frac{B(z^{-1})}{A(z^{-1})} \quad \frac{h_k^{(G)} - h_k}{Q_k} + \epsilon_k \qquad [11]$$

where the constants s and f have been absorbed into $B(z^{-1})$ and h_k has been adjusted by groundwater height $h_k^{(G)}$ to allow for the change of coordinate system mentioned previously (Dietrich, 1988). To predict downstream concentration merely requires addition of upstream concentration $c_0(k)$ to the prediction for y_k which is the salt concentration added within the reach. We restrict attention therefore to prediction of y_k.

Equation 11 is now in the form of equation 2 with nonlinear input $u_k = (h_k^{(G)} - h_k)/Q_k$. It can be solved by the method of refined instrumental variables (Young 1984) which yields asymptotically-efficient estimates under conditions (A1) to (A6) above. Identification of model orders (m,n) can be undertaken using the step-wise strategy given in Young et al. (1980). This yields m=0, n=1 for this case. The estimated model is

$$y_k = - a_1 y_{k-1} + b_0 \frac{h_k^{(G)} - h_k}{Q_k} + \epsilon_k \qquad [12]$$

$$\epsilon_k = -c_1 \epsilon_{k-1} + e_k + d_1 e_{k-1}$$

with mean values of the parameters and their standard errors given in table 1.

Figure 1 shows the model fit to the calibration data set. The model explains about 97 per cent of the output concentration difference in a mean square sense, that is $R^2 = 0.97$. Diagnostic checks show: that the model residuals are not significantly correlated with the input u_k that the noise model estimates of e_k are not significantly autocorrelated and the Durbin-Watson statistic is 2.02; and that the time-varying parameters estimates do not vary within more than ± 5 per cent from their average values. The calibrated model also explains well the salt concentration added within the reach for data sets from three different time periods available. The R^2 values are 0.86, 0.87 and 0.91.

Table 1.
Estimated parameter means and standard deviations for model equation 12.

Parameter	Mean value	Standard Deviation
a_1	-0.794	0.054
b_0	0.105	0.026
c_1	-0.725	0.043
d_1	-0.111	0.061

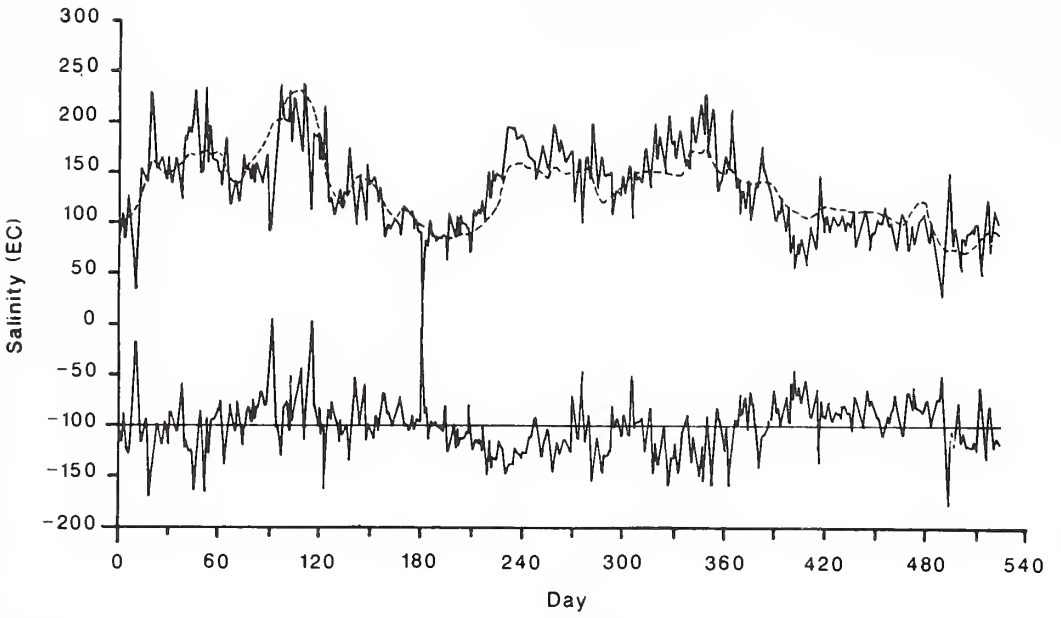


Figure 1.
The model fit (dashed smooth line) applying the parameter values in table 1 to daily saline accession data (continuous line). Model residuals are plotted below. The period of model calibration is from December 5, 1981 to May 12, 1983.

APPLICATION TO POLLUTION INCIDENT PREDICTION

Finally, this section describes briefly the Aggregated Dead Zone (ADZ) model for dispersions in natural channels as an interesting application of a non-linear predictive model based on an application of linear systems analysis to a water quality problem (see Young 1982, Beer and Young 1983, Young and Wallis 1986, Wallis et al. 1988). Many studies have shown that the traditional description of dispersion processes in channels, the advection-dispersion equation (ADE), is not a good descriptor of dispersion in many natural channels where experimental measurements demonstrate the concentration curves have much longer tails than would be predicted by the ADE model (see for example, Day 1975, Sabol and Nordin 1978, Beltaos 1980, Nordin and Troutman 1980, Bencala and Walters 1983, Chatwin and Allen 1985). This has been ascribed to the effects of zones of inefficient mixing or "dead zones" within the flow, probably associated with internal shear produced by the irregularities of natural channels and exchanged with storage in the sediments of the bed and banks which have been shown to be large in some cases (e.g. Zellweger et al. 1986).

The ADZ model is based on the concept that the dominant control on dispersion in natural river channels is inefficient mixing between the mainstream flow and lateral zones of relatively slow moving water. Transport within the mainstream flow is treated as pure advection characterized by a time delay that may vary with discharge. An approximate discrete time mass balance equation for a conservative solute in a reach containing a single effective lateral mixing volume, V_e, under conditions of steady discharge, Q, may be written, under the assumption that input concentrations are constant over a time interval as

$$c(k+1) = -\exp\left[-\pi \frac{Q}{V_e}\right] c(k) + \left\{1 - \exp\left[-\pi \frac{Q}{V_e}\right]\right\} c_i(k-\tau) \qquad [13]$$

where $c(t)$ = the output concentration from the reach, $[ML^{-3}]$
$c_i(t)$ = input concentration to the reach at time t, $[ML^{-3}]$
π = the length of the time step [T] and
τ = the time delay [T].

The mean residence time of the solute within the mixing volume is then given by $T = V_e/Q$.

Equation 13 is the equation of the simplest first order ADZ mixing element. The mean travel time of a conservative solute within the reach is the sum of the time delay, τ, and the mean residence time, T, within the mixing volume. We may then write

$$Q = V\,(T + \tau)$$

where V is the total volume within the reach. Dispersion within the element is controlled by the "dispersive fraction" parameter V_e/V, which is equivalent to $T/(T + \tau)$.

Equation 13 has the form of equation 2a with

$$a_1 = \exp\left[-\pi \quad \frac{Q}{V_e} \right] \text{ and } b_0 = (1 + a).$$

Combining a number of such first order model elements in series and/or parallel results in higher order linear ADZ models in the form of equation 2. These models remain linear in their parameters and are amenable to statistically efficient recursive estimation of the parameters using refined instrumental variables (Young 1984) as before. Studies of a variety of reaches have shown that the ADZ model provides very good fits to observed tracer concentration data (figure 2) in comparison with the traditional ADE approach. A major advantage of the ADZ approach is that it makes few a priori assumptions about the nature of the dispersion process in natural channels, but can be used to determine a complexity of model structure that is justified by the experimental data (see Wallis et al. 1988). We show below how the ADZ approach can also be used as a predictive tool. Results of fitting the ADZ model to a wide variety of reaches have shown that first order models may be appropriate to model the transport of an already dispersed concentration wave through reaches up to tens of kilometers long.

Equation 13 applies for the case of steady-state flows, but we know that the dispersion characteristics of a given river reach will vary with changes in discharge. Detailed studies over a variety of discharges for several reaches have shown that, compared with the dispersion coefficient of the ADE model, the dispersive fraction is a very conservative parameter and to a first approximation can be assumed to be a constant for a reach (figure 3a, Wallis et al. 1988). To predict the dispersion at any steady discharge it will then be necessary to estimate this dispersive fraction and the nonlinear change in mean travel time $\mu = (T + \tau)$ with discharge.

Given a small sample of tracer measurement for a reach, it will often be possible to fit a simple model for the variation in travel time with discharge, such as

$$\mu = f + g\beta \qquad [14a]$$

where $\beta = 1/Q$. This may be posed as a linear regression problem in the variables β and μ. Figure 3b shows such a regression fitted to four of the available measurement points for a short reach of the River Brock in Lancashire, UK. Prediction for other discharges will result in an uncertain estimate of the mean travel time with a standard error that is given by

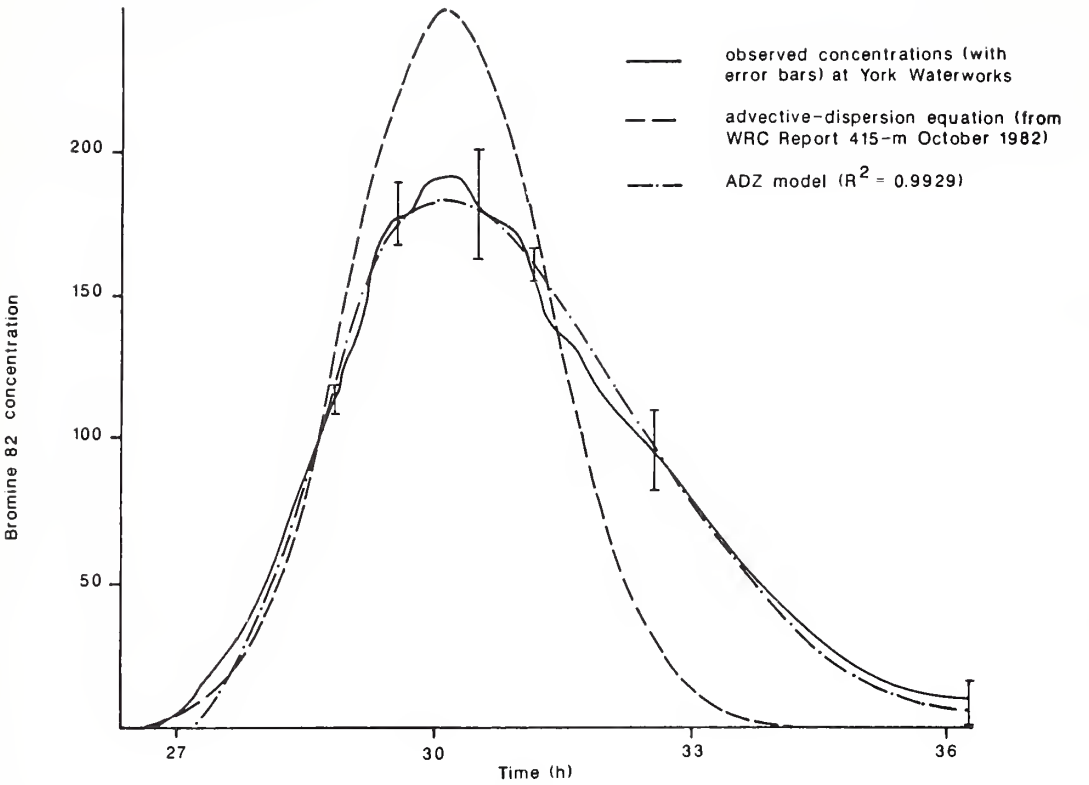


Figure 2.
Observed and predicted concentrations using both the advective-dispersive equation model and ADZ model for a tracer experiment on the River Ouse, Yorkshire, England. Measurement data provided by Yorkshire Water and Water Research Centre.

$$\sigma_\mu^2 = \frac{n}{n-2}(1 - r^2)\, s_\mu^2 \left[1 + \frac{(\beta - \bar{\beta})^2}{s_\beta^2} \right] \qquad [14b]$$

where s^2 = the sample variance of the n samples of either μ or $\beta = 1/Q$, and
r = the regression correlation coefficient (see for example Benjamin and Cornell 1970, p. 435).

Under the assumption that the dispersive fraction has a constant mean and is independent of mean travel time, then a standard error can also be estimated from the available data. We then have all the information required to set up a first order ADZ model of dispersion in the reach, together with estimates of the uncertainty in the parameters.

Estimation of the uncertainty in the predictions is, however, not a simple linear problem because of the dependence of the dispersive fraction on the aggregate sum of T and τ. Quantitative prediction using equation 13 requires disaggregation of the time delay. However, for this simple model implementation of a Monte Carlo procedure for assessing predictive uncertainty is straightforward as follows. Given the discharge for which predictions are required, the mean and standard error of the mean travel time are predicted using equation 14. The i th realization of the first order parameters is then obtained from

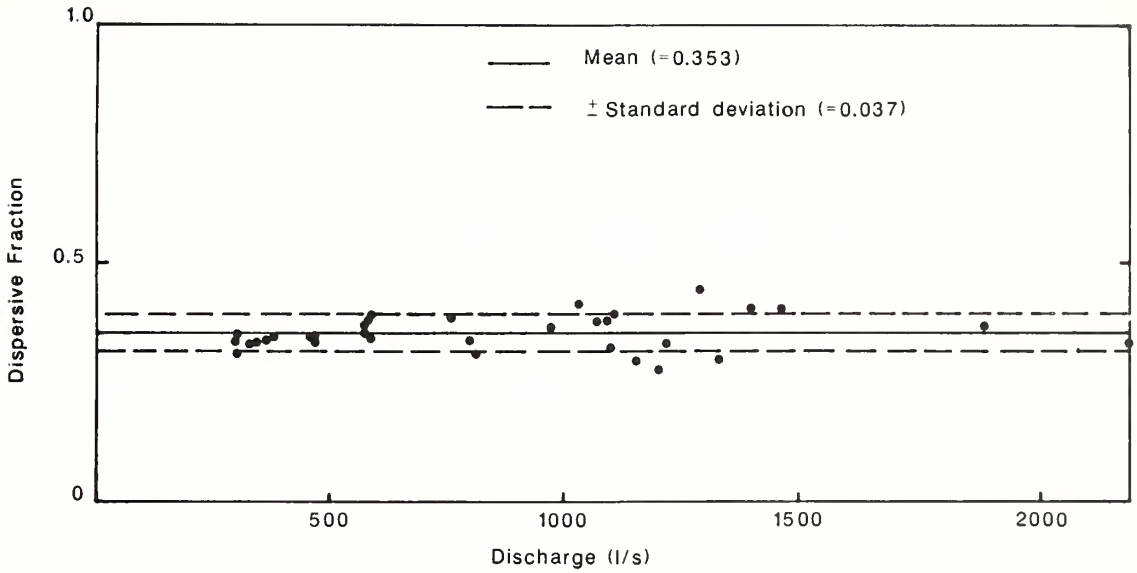


Figure 3a.
The variation in dispersive fraction with discharge for tracer measurements on a 150-m reach of the River Brock, Lancashire, England.

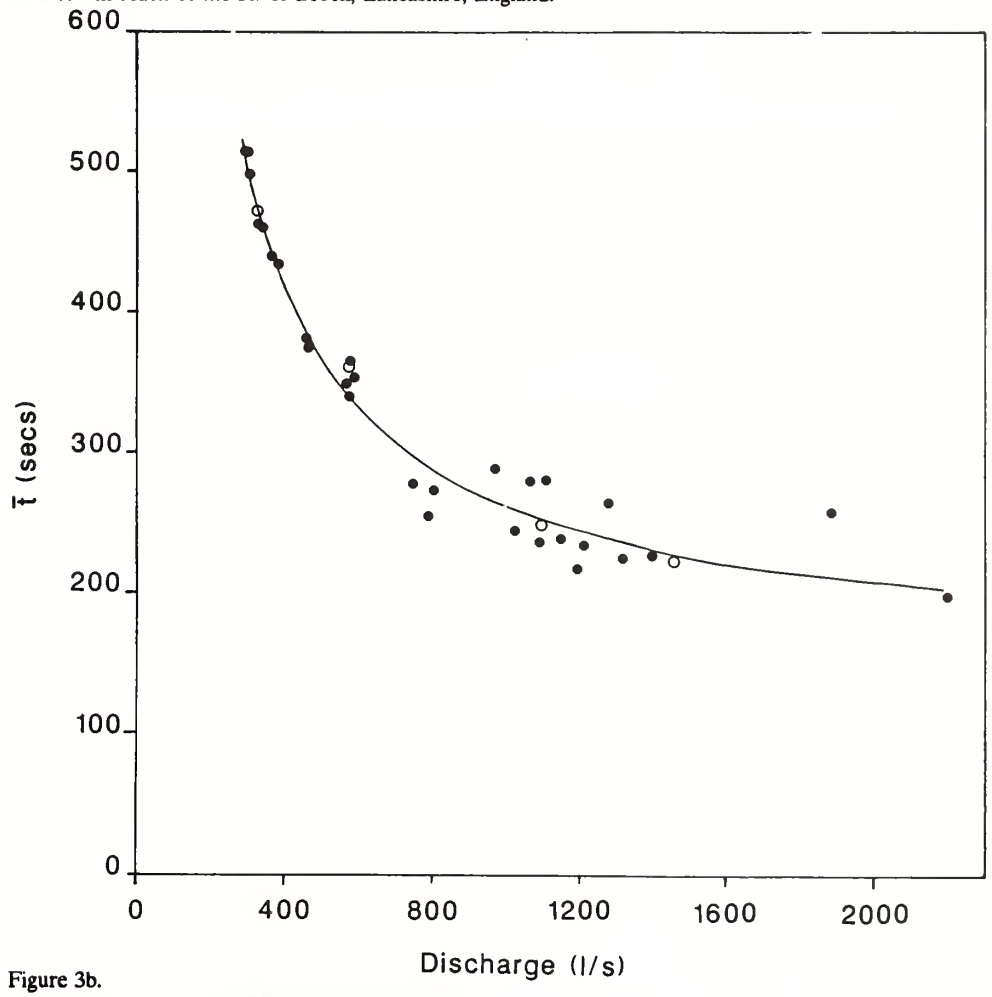


Figure 3b.
The variation in mean travel time with discharge for the River Brock measurement reach. Large open symbols indicate points used in regression for predictive model.

$$\mu_i = f + g\beta + \epsilon_1 \sigma_\mu \; ; \; (V_e/V)_i = \overline{(V_e/V)} + \epsilon_2 \sigma V_e/V \qquad [15a]$$

$$T_i = \mu_i (V_e/V)_i \; ; \; \tau_i = \mu_i - T_i \qquad [15b]$$

$$a_i = \exp(-\pi/T_i) \; ; \; b_0 = 1 + a_1 \qquad [15c]$$

where ϵ_1 and ϵ_2 are random variables chosen from the normal distribution with mean zero and variance one.

The ADZ model is run for many such realizations and the mean and variance of the predicted concentrations calculated for each time step. Tests showed that 100 realizations were sufficient for convergence. Figure 4 shows the results of a comparison of predictions made for a number of different discharges on the basis of the four measurements of figure 3 with observed tracer curves for those discharges.

SUMMARY

This paper has outlined some rational approaches to the problems of predicting the responses of complex real-world systems about which we have only limited information. If is pointed out that although there is a temptation to build in as much physical information as possible about the system, the seductiveness of overly-complex models should be resisted, since the parameters of such models cannot generally be adequately identified, nor can their uncertainties always be quantified, given the limited information available. Approaches to reducing the dimensionality of

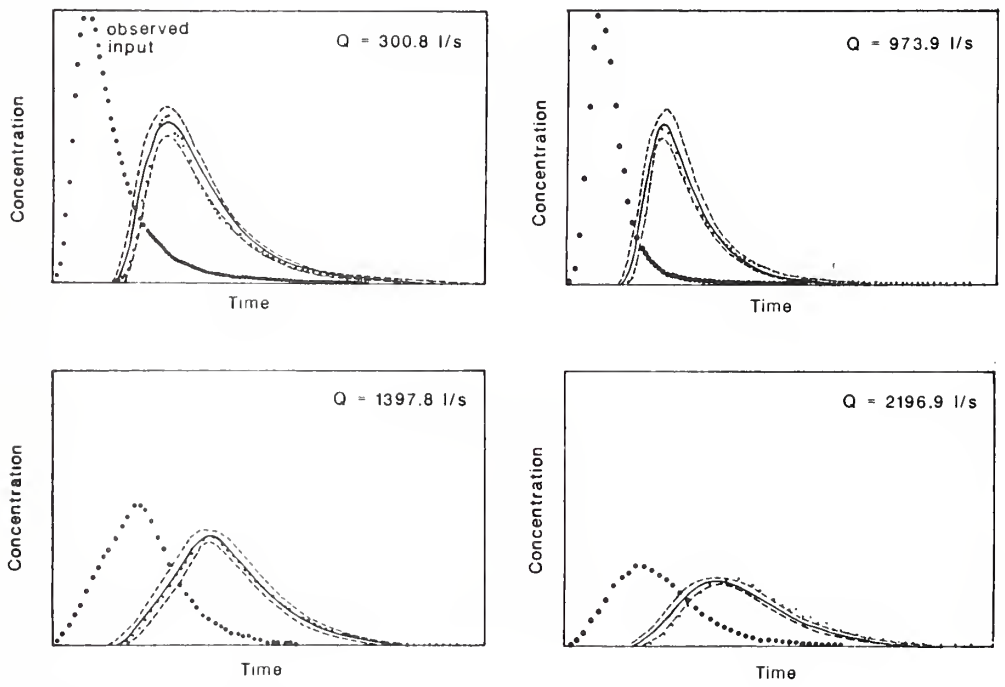


Figure 4.
Predicted tracer concentration curves at different discharges on the River Brock obtained using a first order ADZ model. Solid lines represent mean prediction of 100 Monte Carlo runs, dotted lines represent two standard deviations about the mean. Point symbols represent measured concentrations but these have not been used in obtaining the predictions which are based only on the four measurements indicated in figure 3.

the model are considered, including sensitivity analyses and the use of stochastic models. This leads to a discussion of the sources of predictive uncertainty in models of complex systems. It is argued that all predictive models should attempt to assess the uncertainty associated with their predictions. Techniques for estimating predictive uncertainty are discussed. Finally, two example studies are given to demonstrate different ways in which physical knowledge about system behavior may be combined with stochastic modeling techniques to obtain predictions, together with approximate uncertainty estimates. The first is a model of salinity management for the River Murray in Australia taking account of interactions between surface and groundwaters. The second uses the Aggregated Dead Zone model to predict the dispersion of a pollutant in a river at any discharge, extrapolating the behavior observed during a small number of tracer tests.

REFERENCES

Anderson, O.D., and M.R. Perryman (eds). 1982. Applied Time Series Analysis. North-Holland, Amsterdam.

Anderssen, R.S., and F.R. de Hoog. 1983. The nature of numerical processes, Mathematical Scientist. 8:115-141.

Bear, J. 1972. Dynamics of fluids in porous media. Elsevier, New York.

Beck, M.B. 1983. Uncertainty, system identification and the prediction of water quality, In Uncertainty and Forecasting of Water Quality, M.B. Beck and G. van Stratten (eds), 3-68, Springer-Verlag, New York.

Beer, T., and P.C. Young. 1983. Longitudinal dispersion in natural streams. J. Environ. Eng. ASCE, 109:10349-1067.

Beltaos, S. 1980. Longitudinal dispersion in rivers. J. Hydraul. Div. ASCE. 106:151-172.

Bencala, K.E., and R.A. Walters. 1983. Simulation of solute transport in a mountain pool and riffle stream: A transient storage model. Water Resour. 19:718-724.

Benjamin, J.R., and C.A. Cornell. 1970. Probability, Statistics and Decision for Civil Engineers. McGraw Hill, New York.

Bennett, R.J, and R.J. Chorley. 1978. Environmental systems: Philosophy, analysis and control. Methuen, London.

Beven, K.J. 1979. A sensitivity analysis of Penman-Monteith actual evapotranspiration estimates. J. Hydrol. 44:169-190.

Beven, K.J. 1987. Towards a new paradigm in hydrology, in Water for the Future: Hydrology in perspective. IAHS Publ 164:393-403.

Beven, K.J. 1989. Changing ideas in hydrology: the case of physically-based models. J. Hydrol. 105:157-172.

Box, G.E.P., and G.M. Jenkins. 1976. Time series analysis: Forecasting and control. Holden-Day, San Francisco.

Chatwin, P.C., and C.M. Allen. 1985. Mathematical models of dispersion in rivers and estuaries. Ann. Rev. Fluid Mech. 17:119-149.

Day, T.J. 1975. Longitudinal dispersion in natural channels. Water Resour. Res. 11:909-918.

Dietrich, C.R. 1988. Inverse problems in surface and subsurface hydrology. PhD Thesis, Australian National University.

Dietrich, C.R., A.J. Jakeman and F. Ghassemi. 1988. Modeling for surface and subsurface hydrological management. Proc. IMACS 12th World Congress on Scientific Computation, Volume 2, R. Vichnevestsky et al. (eds.), Paris, July 18-22, pp. 396-398.

Dietrich, C.R., A.J. Jakeman and G.A. Thomas. 1986. Statistical models for solute travel time under unsteady flow conditions. J. Hydrol. 88:253-2734.

Dietrich, C.R., A.J. Jakeman and G.A. Thomas. 1988. Prediction of solute concentration in streams subject to groundwater inflow. Institution of Engineers Australia, National Conference Publication 88/1, 75-79.

Efron, B., and R. Tibshirani. 1986. Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy. Stat. Sci. 1:54-77.

Frank, P.M. 1978. Introduction to system sensitivity theory. Academic, New York.

Gardner, R.H., D.D. Huff, R.V. O'Neill and others. 1980. Application of error analysis to a marsh hydrology model. Water Resour. Res. 16:201-209.

Garen, D.G., and S.J. Burgess. 1981. Approximate error bounds for simulated hydrographs. J. Hydraul. Div. ASCE. 107:1519-1534.

Guymon, G.L., M.E. Harr, R.L. Berg and T.V. Hromadka II. 1981. A probabilistic-deterministic analysis of one-dimensional ice segregation in a freezing soil column. Cold Reg. Sci. Tech. 5:127-140.

Hornberger, G.M., and R.C. Spear. 1981. An approach to the preliminary analysis of environmental systems. J. Environ. Manage. 12:7-18.

Hornberger, G.M., K.J. Beven, B.J. Cosby and D.E. Sappington. 1985. Shenandoah watershed study: Calibration of a topography based, variable contributing area hydrological model to a small forested catchment. Water Resour. Res. 21:1841-1850.

Hornberger, G.M., B.J. Cosby and J.N. Galloway. 1986. Modeling the effects of acid deposition: Uncertainty and spatial variability in estimation of long term sulfate dynamics in a region. Wat. Resour. Research. 22:1293-1302.

Jakeman, A.J. 1979. Multivariable instrumental variable estimators: The choice between alternatives. Proc. 5th IFAC Symposium on Identification and System Parameter Estimation. R. Isermann, Ed., 1:459-467, Pergamon Press.

Jakeman, A.J. 1985. Application of identification and system parameter estimation to environmental problems: Some recent examples. Proc. 7th IFAC Symposium on Identification and System Parameter Estimation. H.A. Barker and P.C. Young, Eds. 1:445-450, Pergamon Press.

Jakeman, A.J. 1988. Features and principles of a systems approach to environmental modeling. Proc. 12th IMACS World Congress of Scientific Computation, Volume 2, R. Vichnevetsky et al. (eds.), Paris, July 18-22, pp. 380-385.

Jakeman, A.J., and P.C. Young. 1979. Joint parameter/state estimation. Electron. lett. 15:582-583.

Jakeman, A.J., and P.C. Young. 1983. Advanced methods of recursive time series analysis. Int. J. Control. 37:1291-1310.

Jakeman, A.J., and P.C. Young. 1984. Recursive filtering and smoothing procedures for the inversion of ill-posed causal problems. Utilitas Mathematica. 25:351-376.

Jakeman, A.J., M.A. Greenaway and J.N. Jennings. 1984. Time series models for the prediction of stream flow in a karst drainage system. J. Hydrol. (New Zealand). 23:21-33.

Jakeman, A.J., L.P. Steele and P.C. Young. 1980. Instrumental variable algorithms for multiple input systems described by multiple transfer functions. IEEE Transactions of Systems, Man and Cybernetics. 10:593-602.

Mahendrarajah, S., P.C. Young and A.J. Jakeman. 1982. Stochastic time series analysis and the village dam management problem in Sri Lanka, on optimal allocation of water resources. IAHS Publ. no. 135:251-2634.

McCuen, R.H. 1976. The anatomy of the modeling process. In C.A. Brebbia, Ed. Mathematical Models for Environmental Problems. Pentech Press.

Nordin, C.F., and B.M. Troutman. 1980. Longitudinal dispersion in rivers: the persistence of skewness in observed data. Wat. Resour. Res. 16:123-128.

Pierce, D.A. 1972. Least squares estimation in dynamic-disturbance time series models. Biometrika. 59:73-78.

Rabitz, H., M. Kramer and D. Dacol. 1983. Sensitivity analysis in chemical kinetics. Ann. Rev. Phys. Chem. 34:419-61.

Rogers, C.C.M., K.J. Beven, E.M. Morris and M.G. Anderson. 1985. Sensitivity analysis, calibration and predictive uncertainty of the Institute of Hydrology Distributed Model. J. Hydrol. 81:179-187.

Rosenblueth, E. 1975. Point estimates for probability moments. Proc. Nat. Acad. Sci. USA. 72:3812-3814.

Sabol, G.V., and C.F. Nordin. 1978. Dispersion in rivers as related to storage zone. J. Hydraul. Div. ASCE. 104: 695-708.

Schweppe, F.C. 1973. Uncertain dynamic systems. Prentice-Hall, Englewood Cliffs, New Jersey.

Spear, R.C., and G.M. Hornberger. 1980. Eutrophication in Peel Inlet. II. Identification of critical uncertainties via generalized sensitivity analysis. Water Res. 14:43-49.

Taigbenu, A., and J.A. Liggett. 1986. An integral solution for the diffusion-advection equation. Water Resour. Res. 22:1237-1246.

Thomann, R.V. 1972. Systems analysis and water quality management. McGraw-Hill, New York.

Thomas, G.A., and A.J. Jakeman. 1985. Management of salinity in the River Murray Basin. Land Use Policy. 2:87-102.

Wallis, S.G., P.C. Young and K.J. Beven. 1988. Experimental investigation of the aggregated dead zone model for longitudinal solute transport in stream channels, submitted to Instit. Civ. Eng. Proc. [In Press].

Whitehead, P.G., P.C. Young and G.M. Hornberger. 1979. A systems model flow and water quality in the Bedford-Ouse River I. Stream modeling. Water Res. 13:1159-1169.

Whitehead, P.G., and G.M. Hornberger. 1984. Modeling algal growth in the River Thames. Water Res. 18:945-953.

Young, P.C. 1978. A general theory of modeling for badly defined systems, In The Modeling of Environmental Systems. G.C. Vansteenkiste (Ed.), North Holland, Amsterdam. 103-135.

Young, P.C. 1982. The validity and credibility of models for badly defined systems, In Uncertainty and Forecasting of Water Quality. M.B. Beck and G. van Straten, Eds. Pergamon, Oxford.

Young, P.C. 1984. Recursive estimation and time series analysis. Springer-Verlag, Berlin.

Young, P.C., and A.J. Jakeman. 1979. Refined instrumental variable methods of recursive time-series analysis, II. Single input, single output systems. Int. J. Control. 29:1-30.

Young, P.C., A.J. Jakeman and R.E. McMurtrie. 1980. An instrumental variable method for model order identification. Automatica. 16:281-294.

Young, P.C., and S.G. Wallis. 1986. The aggregated dead zone model for dispersion in rivers. Proc. Int. Conf. Water Quality Modeling in the Inland Natural Environment. BHRA.

Zellweger, G.W., V.C. Kennedy, K.E. Bencala and others. 1986. Data on the solute concentration within the subsurface flows of Little Most Man Creek in response to a transport experiment, Redwood National Park, Northwest California. USGS Open File Report 86-403W.

SELECTION, APPLICATION, AND VALIDATION OF ENVIRONMENTAL MODELS

A.S. Donigian, Jr.[1] and P.S.C. Rao[2]

ABSTRACT

Model validation is becoming an increasingly important issue for both scientists and engineers involved in model development, and for managers involved in environmental decision-making. With the increasing use of mathematical models in environmental analyses, questions relative to the proper selection and application of a specific model for a specific environmental setting must be resolved in determining whether or not the 'correct' model was used in an 'appropriate' manner so that the validity of the model results can be assessed. Decision-makers are demanding that these issues be addressed before accepting the model results. Guidelines have been, and are being, developed to help model users in the proper selection and application of models for specific environmental media and contamination problems. However, guidelines for model validation are non-existent and efforts are currently underway to develop them. This is a critical need because model validation is the 'bottom line' in determining the acceptance and use of models in environmental decision-making.

INTRODUCTION

In recent years there has been a dramatic growth in the number of mathematical models and in their use to estimate environmental concentrations of chemicals to which humans and other organisms may be exposed. The development of environmental models in the late 60's and throughout the 70's has led to the application of modeling for the analysis of a wide range of critical regulatory issues pertaining to environmental contamination. Although model development and refinement are continuing in the 80's, there has been a clear shift in emphasis from development to application of these models to 'solve' specific environmental problems.

This shift has been accompanied by the phenomenal growth in computing speed, power, and accessibility that has led to an increasing disparity in our ability to formulate and solve increasingly complex simulation models, our ability to truly understand the environmental processes we are attempting to represent, and our ability to provide values for a wide variety of required model input data. This imbalance has focused concern on issues related to the proper selection of models for specific environmental problems, the proper application procedures needed for these models, and the degree of validation required to instill confidence in the model predictions. This paper explores these issues of model selection, application, and validation. Emphasis will be on the model validation process because of its importance and its relatively primitive stage of development when compared to selection and application of models. Recent efforts in developing model validation procedures and associated model performance criteria are reviewed in an attempt to identify where we are and where we need to go in obtaining a consensus of what constitutes model validation.

[1]A.S. Donigian, Jr., President and Principal Engineer, AQUA TERRA Consultants, Mountain View, CA.

[2]P.S.C. Rao, Professor, Soil Science Department, University of Florida, Gainesville, FL.

FRAMEWORK FOR DISCUSSION

Part of the confusion surrounding model selection, application, and validation is due to the different meanings that have been associated with the various terms - application, calibration, verification, validation, testing - used in the technical literature. Recently, the American Society for Testing and Materials (ASTM, 1984) prepared a Standard Protocol for evaluating environmental chemical-fate models, including proposed standard definitions for selected modeling terms (shown in table 1). Although not all modelers, or model users, will agree completely with these specific definitions, the standardization they provide may help to facilitate communications among modelers and with other professionals. In addition, there are indications that these definitions, or slight modifications, are gaining acceptance by other groups (discussed below).

In this paper, we adhere to the ASTM definitions. We view model selection and model application as two separate, but complementary processes in the use of models for environmental analyses. Clearly, a model must be selected for a specific application, based on the environmental system characteristics, the questions to be resolved, the available resources, etc. before it can be applied.

After a model has been selected for the application, a series of steps or procedures are followed that comprise the modeling, or model application, process. These steps, as listed in table 2, can be grouped into three phases: Phase I includes data collection, model input preparation, and

Table 1.
Definition of terms used in ASTM standard practice for evaluating environmental fate models of chemicals.

Algorithm -	the numerical technique embodied in the computer code.
Calibration -	a test of a model with known input and output information that is used to adjust or estimate factors for which data are not available.
Compartmentalization -	division of the environment into discrete locations in time or space.
Computer Code (computer program) -	the assembly of numerical techniques, bookkeeping, and control language that represents the model from acceptance of input data and instructions to delivery of output.
Model -	an assembly of concepts in the form of a mathematical equation that portrays understanding of a natural phenomenon.
Sensitivity -	the degree to which the model result is affected by changes in a selected input parameter.
Validation -	comparison of model results with numerical data independently derived from experiments or observations of the environment.
Verification -	examination of the numerical technique in the computer code to ascertain that it truly represents the conceptual model and that there are no inherent numerical problems with obtaining a solution.

Source: ASTM (1984)

parameter evaluation; Phase II is model testing; and Phase III is analysis of alternatives. Thus we view model testing as part of the model application process.

Before discussing model selection, we will provide an overview of the various steps in the modeling process because it establishes an overall framework for definition of terms and subsequent sections on model selection, application and validation.

Overview of the Modeling Process

Phase I provides the groundwork for the modeling effort by developing, collecting, discovering, and preparing the data needed for model application. This includes the observed meteorologic, hydrologic and water quality data in addition to site characterization information. Data collection may also include sample collection and analysis if adequate historical data are not available. Model input preparation is the process of preparing the collected data in a format acceptable to the model. Parameter evaluation is the process of estimating the specific model parameters required by the model, based on site characteristics, evaluation guidelines, and prior experience. At the conclusion of Phase I, model execution runs can be initiated.

As shown in table 2, model testing is Phase II of the modeling process. As noted above, part of the confusion surrounding the model testing phase is largely because different meanings have been attached to the terms calibration, verification, validation, and post-audit in the technical literature. The process of model testing should ideally include all three steps shown in table 2. We say "ideally" because in many applications existing data will not support performance of all steps. In chemical-fate modeling, measured data for validation are often lacking and post-audit analyses are rare for any type of modeling exercise.

Calibration

Calibration is probably the most misunderstood of all the model validation components. Calibration is the process of adjusting selected model parameters within an expected range until the differences between model predictions and field observations are within selected criteria for performance. For all operational, deterministic models (or portions thereof), calibration is usually needed and highly recommended. Calibration is needed to account for spatial variations not represented by the model formulation; functional dependencies of parameters that are either non-qualifiable, unknown, and/or not included in the model algorithms; or extrapolation of laboratory measurements of parameters to field conditions. It is clear that the need for calibration

Table 2.
The modeling process.

Phase I	• Data Collection • Model Input Preparation • Parameter Evaluation	
Phase II	• Calibration • Validation • (Post-Audit)	Model Testing
Phase III	• Analysis of Alternatives	

increases the user effort and data required to appropriately apply a model. However, any model can be operated without calibration depending on the extent to which critical model parameters (usually refined through calibration) can be estimated from past experience and other data. In the area of modeling pesticide runoff, Lorber and Mulkey (1982) have shown that so-called 'calibration-independent' (empirical, in this case) models produced their best results....."only after some deliberation and reassignment of initial parameter estimates." This in effect was a calibration process.

Validation

Validation is the complement of calibration; model predictions are compared to field observations that were not used in model development or calibration. This is usually the second half of split-sample testing procedures, where the universe of data is divided (either in space or time), with a portion of the data used for calibration and the remainder used for validation. In essence, validation is an independent test of how well the model (with its calibrated parameters) is representing the important processes occurring in the natural system. Although field and environmental conditions are often different during the validation step, parameters determined during calibration are not adjusted during validation.

Validation and verification have been used interchangeably by many investigators. Validation, in essence, means determining the model's ability, after calibration, to represent a specific site and/or a specific model application; whereas the ASTM definition of verification is restricted to verifying the operation of the numerical procedures in the code.

Post-Audit Analyses

Post-Audit Analyses are the ultimate tests of a model's predictive capabilities. Model predictions for a proposed alternative are compared to field observations following implementation of the alternatives. The degree to which agreement is obtained based upon the acceptance criteria reflects on both the model capabilities and the assumptions made by the user to represent the proposed alternative. Unfortunately, post-audit analyses have been performed in few situations, and thus it is noted in parentheses in table 2.

The final and perhaps most critical phase of the modeling process is the use of a model as a decision aid for assessing management or regulatory alternatives. In analyzing various alternatives, the validated model is used as a tool to predict the changes in system response resulting from a proposed alternative; this alternative may be represented by adjustments (changes) to model input, parameters, and/or system representation. During the model testing phase, the model results are compared with observed data for selected time periods; whereas, in the analysis of alternatives the model results for a specific alternative are compared to model results produced by appropriate base conditions. In this way the relative changes in system response associated with a proposed alternative can be identified and analyzed.

Before management alternatives can be analyzed, the model testing phase must proceed to the point where model results are sufficient to demonstrate that the model provides a realistic and credible representation of system behavior.

MODEL SELECTION

There are literally hundreds of environmental models available, many of which may be appropriate for application to a specific problem depending on characterization of the pollutant sources (e.g., point versus nonpoint), the type of environmental system (e.g., surface water versus groundwater versus estuarine/marine environments), and the level of analysis required (e.g., screening-level versus site-specific). Our discussion thus far has been generic and rather generalized because modeling is used in support of a wide variety of environmental issues, ranging from the evaluation of pollutant concentrations resulting from ocean disposal, to the characterization of fate/transport

profiles of potential contaminants of concern, to the analysis of the relative impacts of landfilling versus land application of sludge. For each of these types of problems, and many others, environmental models may need to be applied in a site-specific mode with a detailed characterization of the project site, or in a regional or national mode with proper selection of "representative" input data.

Because of the diversity of pollutant sources, environmental systems, and levels of analysis that may be required, and the large number of available models, a generalized framework for model selection is needed to insure use of models appropriate for the specific problem and to establish consistency in modeling procedures. Such a generalized framework has been developed for EPA for selection and use of models for evaluating the effectiveness of remedial actions at uncontrolled hazardous waste sites (Boutwell et al. 1985). Although developed under Superfund (major U.S. program for decontaminating hazardous waste sites), the framework is sufficiently general that it provides a generic set of procedures (i.e., flow charts and matrices) to guide model selection for most surface and soils/groundwater modeling problems. In fact, the specific models discussed and described in the report (referred to as the "Remedial Action Modeling" report) were originally developed for a wide variety of environmental conditions and contaminants, not specifically for hazardous wastes.

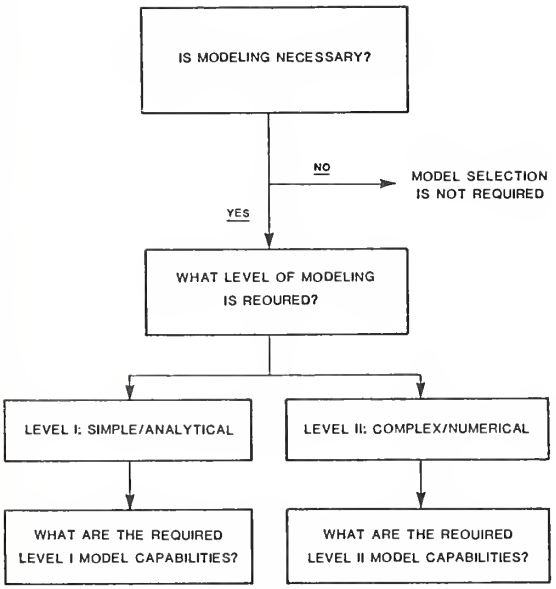
Figure 1 contains the first two flowcharts included in the Remedial Action Modeling report that provide an overview of the model selection process and address the basic issue - "Is modeling necessary?". In many cases, the available data and/or the demands of the analysis may not require a modeling approach. The report describes two modeling categories: Level I refers to simple/analytical models or procedures while Level II includes complex/numerical models. The flowcharts for choosing which level of modeling is required are shown separately in Figure 2 for surface water and soils/groundwater systems. As indicated in the flowcharts, the choice of model level depends on the accuracy requirements of the analysis, the characterization of the system and contaminants, and the available resources. Although the specific questions may differ for surface waters or soils/groundwater systems, and for specific problem applications the general framework for selecting the appropriate modeling level is widely applicable.

Once the level of modeling is selected, the required model capabilities and associated criteria must be established so that a specific model or set of models can be selected. Figure 3 shows the decision flowcharts for defining the required model capabilities for surface and soils/groundwater systems. For numerical models, the selection criteria depend on the processes that must be represented, dimensionality assumptions, time frame (e.g., storm event versus mean annual), and the available data and resources. Since simplified/analytical models have serious limitations in adequately representing management practices or remedial actions, model selection depends on whether analytical models are available for the specific conditions or practices to be evaluated.

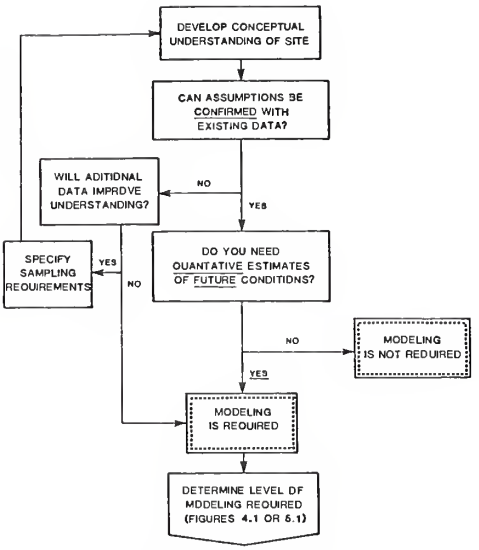
The U.S. EPA Exposure Assessment Group, within the Office of Health and Environmental Assessment, has recently proposed similar model selection criteria and procedures for surface water (U.S. EPA, 1987a), groundwater (U.S. EPA, 1987b), and air (U.S. EPA, in preparation) to provide technical guidance to model users and advance the art of exposure/risk assessment. The selection criteria are supplemented with summary model descriptions to assist in the choice of a specific model appropriate for the user and the problem assessment.

Available Models and Methodologies

Once the model selection criteria and required capabilities have been established, this information is then compared to the characteristics and capabilities of the available models. Considering the hundreds of models available, this could be a significant, if not an enormous effort; fortunately, a variety of model reviews and comparisons have been performed and summarized as valuable sources of information on model characteristics/capabilities. Donigian and Beyerlein (1985)

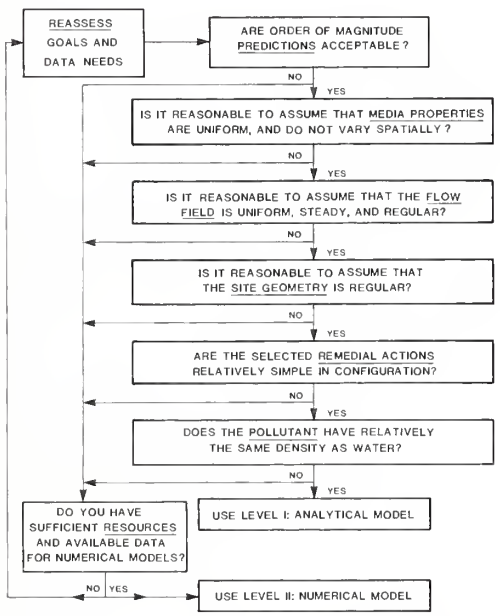


Three basic decisions in model selection.

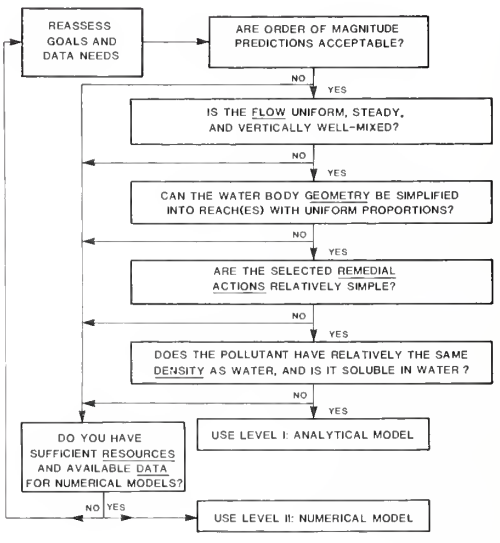


Flow chart to determine if modeling is required.

Figure 1.
Decision flowcharts for model selection and need for modeling (Boutwell et al. 1985).



Flow chart to determine the level of modeling required for soil and groundwater systems.



Flow chart to determine the level of modeling required for surface water systems.

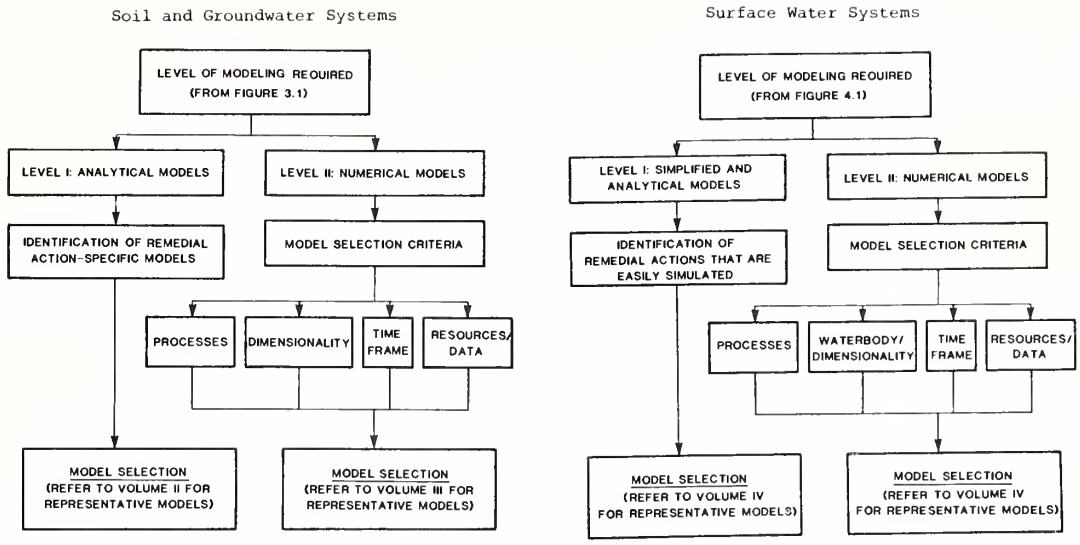


Figure 3.
Decision flowcharts for required model capabilities (Boutwell et al. 1985).

recently reviewed nonpoint source runoff and integrated watershed models was performed by their work, summarized in table 3 emphasized runoff and integrated (i.e., both runoff and instream processes) water quality models that could be used for BMP (Best Management Practice) evaluation for control of nonpoint sources. In a companion effort, HydroQual (1985) reviewed available receiving water models that could be used to accept nonpoint source loadings and evaluate their effects. Boutwell et al. (1985) reviewed the capabilities of selected toxic chemical fate and transport models as part of their model selection and application process for chemical spills in surface waters; a sample of their results is shown in table 4.

For soils and groundwater systems, a recent review is included in the Remedial Action Modeling report. Table 5 summarizes the capabilities of the models reviewed by Boutwell et al. (1985) for representing the surface, unsaturated, and saturated (i.e., groundwater) zones.

These charts and tables are simply examples of the type of information available for model selection. Many other specialized reviews have been performed, such as for vadose zone models of organics (Hern and Melancon 1986), synfuel risk assessments (Donigian and Brown 1983), and geohydrochemical models for the electric power industry (Kincaid et al. 1984). All this information, and additional data from other model reviews, can be used in conjunction with the defined model criteria and required capabilities to insure that appropriate models are selected to meet the needs of any water quality assessment.

MODEL APPLICATION

After selection of the appropriate model(s), the model application process, as described above and in table 2, must be carefully executed and monitored to ensure that the model is operated and applied correctly. There are both technical and operational issues in model application that are critical to the success of the modeling study. The technical issues are primarily related to the accuracy and/or appropriateness of the input data, and to the linkage of models when more than one model is required to be used conjunctively. Models can be applied either to a specific site or

Table 3.
Nonpoint source and integrated watershed models
(Donigian and Beyerlein 1985).

CHARACTERISTICS AND CAPABILITIES OF SELECTED NPS RUNOFF PROCEDURES AND MODELS

	LAND USE/LOAD SOURCES						HYDROLOGY			WATER QUALITY			TIME SCALE			DATA NEEDS		SPACE SCALE		
	Urban	Agriculture	Forest/Natural	Mining	Precipitation	Chemical Application	Surface Runoff	Subsurface Flow	Snowmelt	Sediments	Nutrients	Pesticides/Toxics	Annual Loads	Event Loads	Continuous Simulation	Detailed	Moderate	Minimal	Segmented/Multiple Catchments	Lumped/Single Catchment	Use, Documentation, Support
LOADING/SCREENING PROCEDURES																					
Hydroscience	●						●		●		●		●					●		●	A/M
EPA Screening Procedures	●	●	●		●					●	●	●	●	●				●		●	A
WRENS			●				●		●	●	●	●	●					●	●		A
WLFNPS		●	●				●		●	●	●			●			●		●		M/A
SWMM - Level I	●						●	○		●	●		●				●			●	A
RUNOFF MODELS																					
Simplified SWMM	●						●				●			●	●		●		●		M
ARM		●				●	●	●	●	●	●	●		●	●	●				●	E/A
NPS	●	●	●				●	●	●	●	●			●	●	●			●		E/A
HSPF/PERLND & IMPLND	●	●	●	●	○	●	●	●	●	●	●	●		●	●	●			●		E
CREAMS/CREAMS 2		●	●			●	●			●	●	●		●	●	●				●	E
ANSWERS		●	●	●			●	●		●	●			●		●			●		A
ACTMO		●				●	●	●	●	●	●	●		●	●	●			●		M
SWMM	●						●	○	●	●	●			●	●	●			●		E
STORM	●						●	○	●	●	●				●		●			●	E
MUNP	●						●			●	●			●			●			●	M
ILLUDAS/DRAINQUAL	●						●	●			●			●		●	●		●		A
DR3M	●				●		●	○		●					●	●			●		A
PRMS			●				●	●	●	●				●	●	●			●		A

Notes: ● - Capability included in model
○ - Capability not explicitly included but can be user-defined

Use/Documentation/Support
E - Extensive
A - Adequate
M - Minimal

CHARACTERISTICS AND CAPABILITIES OF INTEGRATED WATERSHED MODELS

	WATERBODY & FLOW CONDITIONS						WATER QUALITY									TIME SCALE	DATA NEEDS		SPACE SCALE			
	Rivers/Streams	Lakes/Impoundments	Estuaries	Confined Flow	Drainage/Control Structures	Point Source Discharges	Temperature	D.O./BOD/NBOD	Suspended Sediment Transport	Sediment Scour/Deposition	Sediment Contaminant Interaction	Nutrient Kinetics	Pesticides/Toxics	Biologic Simulation	Dynamic	Steady-State	Detailed	Moderate	Multi Land Use Multiple Catchments	Single Land Use Single Catchment	Data File Management	Use, Documentation, Support
INTEGRATED WATERSHED MODELS																						
HSPF	●	●		○	●	●	●	●	●	●	●	●	●	●	●		●		●		●	E
SWMM (RECEIV)	●	●	●	●	●	○		●				●	○	●	●		●		●			E
PRS	●	●											●		●		●		●			A
UTM-TOX	●		●		●	●	○		●	●			●		●		●		●			M
SWAM	●	●					●		●	●		●	●		●		●		●			M/A

Notes: ● - Capability included in model
○ - Capability not explicitly included but can be user-defined

Use/Documentation/Support
E - Extensive
A - Adequate
M - Minimal

under conditions representative of a larger region. In a site-specific application, input data such as meteorologic timeseries (e.g., rainfall) and model parameter values (e.g., land slope, infiltration rates) should be available for the site. If data must be estimated or extrapolated from other areas, then the estimation/extrapolation procedures must be well-established and well-documented. The same is true when models are applied under "regional" conditions; the basis for estimating regionally representative or average values for model input must be clearly described and justified. Alternatively, if there is uncertainty in key parameter values, sensitivity analyses should be applied and/or stochastic procedures (e.g., Monte Carlo simulations) may be needed to establish the uncertainty in model output associated with the variability and/or uncertainty in model input.

Table 4.
Simulated processes vs. models matrix for selected surface water chemical fate models.

PROCESSES / SELECTED NUMERICAL MODELS	WATERBODY: ESTUARY	WATERBODY: LAKE	WATERBODY: RIVER	SPATIAL DOMAIN: 1-D	SPATIAL DOMAIN: 2-D	SPATIAL DOMAIN: 3-D	DISPERSION: TURBULENCE	DISPERSION: EMPIRICAL MIXING	TEMPORAL DOMAIN: STEADY STATE	TEMPORAL DOMAIN: TIDALLY AVG.	TEMPORAL DOMAIN: QUASI-DYNAMIC	TEMPORAL DOMAIN: DYNAMIC	SEDIMENTATION: SEDIMENT TRANSPORT	SEDIMENTATION: SUS. SED. SORPTION	SEDIMENTATION: DIRECT BED EXCHANGE	SEDIMENTATION: "ARMORING"	TRANSFORMATION PROCESSES: PHOTOLYSIS	TRANSFORMATION PROCESSES: OXIDATION	TRANSFORMATION PROCESSES: HYDROLYSIS	TRANSFORMATION PROCESSES: VOLITILIZATION	TRANSFORMATION PROCESSES: BIO-DEGRADATION	TRANSFORMATION PROCESSES: BIO-ACCUMULATION	TRANSFORMATION PROCESSES: LUMPED DECAY OR SINGLE POLLUTANT	TRANSFORMATION PROCESSES: DAUGHTER PRODUCTS
TOXIWASP	•	•	•	C	C	C		•				•	•	•	•		•	•	•	•			•	
HSPF		•	•	B								•	•	•	•		•	•	•	•	•		•	•
TODAM	•		•	•				•				•	•	•	•	•	•	•	•	•	•		•	
EXAMS		•	•	C				•	•					•	•		•	•	•	•	•		•	•
SERATRA			•		V		•	•1				•	•	•	•	•	•	•	•	•	•		•	
DEM	•			B				•	•			•		•									•	
FLESCOT	•	•	•			•	•	•				•	•	•	•	•							•	
FETRA	•				L			•				•	•	•	•	•							•	
LEENDERTSE 2D	•				V		•					•												
LEENDERTSE 3D	•					•	•					•												
LARM	•	•	•		V			•															•	

Source: Boutwell et al. 1985

FOOTNOTES:

1. FETRA DOES EMPLOY FIRST-ORDER DECAY FOR A POLLUTANT, HOWEVER, THIS LUMPED PARAMETER MAY BE TOO SIMPLE FOR POLLUTANTS WITH MULTIPLE DEGRADATION PROCESSES.

LEGEND
S = SINGLE ITEM
B = BRANCHING OR NETWORK
L = LATERALLY INTEGRATED
V = VERTICALLY INTEGRATED
C = COMPARTMENTS

In many instances it may be necessary to link two or more models by having one model provide input to other models in order to fully characterize the environmental system and to meet the needs of the assessment. For example, runoff models are commonly linked to receiving-water models to provide the runoff loads from nonpoint sources. Similarly, vadose (unsaturated) zone models are often linked to groundwater models to provide the contaminant load to the aquifer. An example of such a linkage is shown in figure 4 where the PRZM model (Carsel et al. 1984) provides runoff loadings to the EXAMS model (Burns et al. 1982) and groundwater loadings to the AT123D groundwater model (Yeh et al. 1981). This linkage methodology was developed to estimate surface water and groundwater concentrations resulting from land application of municipal sludge (Donigian and Bicknell 1986). Because of differences in characteristic time and space scales between the models, and the environmental compartments they represent, such model linkages must be carefully implemented and clearly described to explain all the assumptions in the modeling analysis. In addition, the methodology shown in figure 4 was applied on a regional basis, requiring justification and explanation of the regional model input values, as mentioned above.

Operational aspects of model application involve a variety of considerations related to the 'nuts-and-bolts' of model use, including correct mathematical operation, data manipulation and input, output analysis, implementation on personal computers (PCs), etc. Although these may seem to be somewhat mundane issues, they can often be critical to the ultimate success of a modeling study. Planning must allow for obtaining and implementing computer codes for the models that may be selected, including execution of standard test runs and comparison with output provided by the developer/distributor in order to insure correct model application. Such procedures are standard for all models distributed by the EPA Center for Water Quality Modeling in Athens, GA, and have been shown to be effective. The Center has participated in the development of the ANNIE software system (Lumb and Kittle 1986) that provides a user-friendly, interactive capability for pre-processing of input and post-processing of results of environmental

Table 5.
General capabilities of selected saturated, surface and unsaturated zone models (Boutwell et al. 1985).

	SURFACE ZONE					UNSATURATED ZONE				SATURATED ZONE				
	LAND SEGMENT (Single / Multiple)	Sorption	Degradation	Erosion	DIMENSION (1,2,3)	Dispersion	Sorption	Degradation	DIMENSION (1,2,3)	Dispersion	Sorption	Degradation	DOCUMENTATION	MAJOR CODE LIMITATIONS
MODEL NAME (References)														
SURFACE ZONE MODEL														
HSPF (EPA)	(M)	X	X	X	L		X	X	L		X	X	C	
SWMM (EPA)	(M)		X	X									C	● Single segment and no sediment-pollutant adsorption
CREAMS (USDA / Corps of Engrs.)	(S)	X	X	X									C	
SEASOIL (A.D. Little, Inc.)	(S)	X	X		1		X	X					C	● Single segment and does not consider erosion
HELP (EPA / Corps of Engrs.)	(S)												C	
UNSATURATED ZONE MODEL														
FEMWATER / FEMWASTE (ORNL)					2	X	X	X					C	
TRUST / MILTRAN (LBL / Battelle)					2		X	X					I	
COLUMN TRANSPORT WITH SORPTION (Kipp, Kenneth L.; England)					1	X	X	X						● 1 dimensional
ODMOD (Argonne National Lab)					1		X							● 1 dimensional, no dispersion and degradation
NMODEL (Univ. of Florida)					1	X								● 1 dimensional, no adsorption and degradation
PERCOL (Battelle)					1			X	I			X		● 1 dimensional, steady state, no dispersion and adsorption
PRZM (EPA / Athens)	(S)	X	X	X	1	X	X	X						● 1 dimensional
SATURATED ZONE MODEL														
FEWA / FEMA (ORNL)									2	X	X	X	C	
SWIFT (Intera)									3	X	X	X	C	
HCTH (Intera)									3	X	X	X	C	
FE3DGW / CFEST (Battelle)									3	X			I	
AT123D (ORNL)									3	X	X	X	C	● Being analytical the model has limited spatial resolution
PLASM (Prickett & Lonnquist)									1	X	X	X	C	
WASTE (Analytical Science Corp.)									1	X	X	X		● 1 dimensional unsteady state or 2 dimensional steady state
GWSIM-II (Texas Dept. of Water Resources)									2	X				● No adsorption and degradation
MOC (Konikow & Bredehoeft, USGS)									2	X	X	X	C	● No adsorption and degradation
GROUNDWATER COMPUTER PACKAGE (Marlon-Lambert, J.; Canada)									2	X	X	X		● No documentation
PATHS (Battelle)									2		X		C	● Being analytical the model has limited spatial resolution
TRANSCOL / FRACSOL (Prickens, J.F.; Canada)									1		X			● 1 dimensional, no dispersion and degradation
GETOUT (Burkholder, et al)									1	X	X	X		● 1 dimensional
NEWSAM (Ledoux, E.; France)									1		X			● 1 dimensional, no dispersion and degradation
VTT (Battelle)									2				C	● 2 dimensional, no pollutant transport sub-module

Legend:
(M) = Multiple Land Segments
(S) = Single Land Segment
X = Considered
C = Complete Documentation
I = Incomplete Documentation or User's Guide
L = Lumped Parameters
? = Unknown

models. These capabilities are especially important for data-intensive models used for comprehensive environmental assessments. ANNIE has been implemented on PCs for a variety of EPA (e.g., HSPF, PRZM, QUAL2) and U.S. Geological Survey (USGS) models.

Analysis/Interpretation of Results

Clearly, the key task in any modeling study is the analysis and interpretation of the model outputs. Since models are simply tools for a quantitative, systematic analysis of specific environmental problems or issues, they do not provide simple YES or NO answers to a manager, a regulator, or decision-maker. Rather, they usually provide detailed information about the expected response of

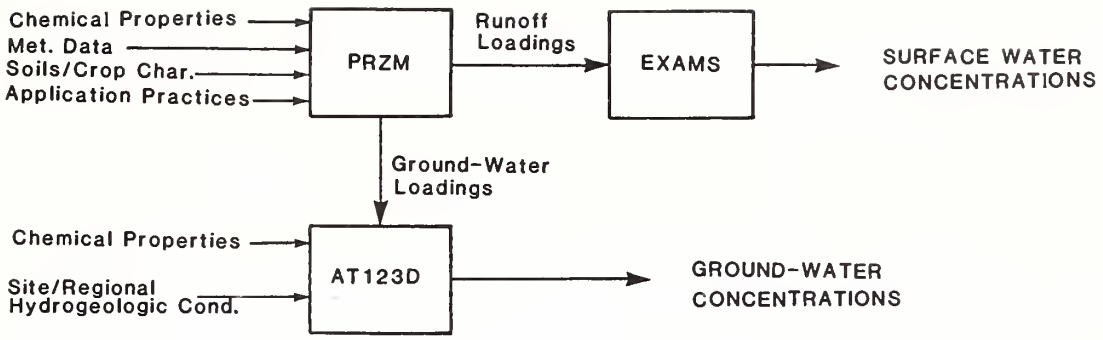


Figure 4.
Model linkage methodology for land application alternative.
(Donigian and Bicknell 1986).

the system to a given perturbation in order that a more informed, objective decision can be made. The mass of computer output generated by models must be analyzed and interpreted in a logical and consistent fashion in order to answer the decision-maker's questions: "What do the results mean?" and, "How accurate and reliable are they?".

In order to understand the true "meaning" of modeling results within a decision- making framework, both the assumptions of the analysis and accuracy expectations (i.e., reliability) must be clearly defined. Both of these considerations are difficult, if not impossible, to discuss in general terms without discussing the specific characteristics of the particular model. However, assumptions usually are included, and required, both in how the model is configured or designed, and how it is applied. Thus, one specific model may be used in many applications, with the same set of model assumptions common to all applications, while the application assumptions may differ from one case to another. Thus, for example, the EXAMS model could be applied at five river sites across the country assuming steady flow conditions at each site, while the specific chemical processes simulated may differ with each chemical analyzed. The decision-maker or analyst must be cognizant of both kinds of assumptions, and their associated limitations, in order to appreciate the validity of the modeling results.

The accuracy associated with the results of modeling studies depends on the specific model used, the accuracy of the input data, the characterization of the environmental system being simulated, and the expertise/experience and resources available to the model user. Decision-makers must understand that all these factors determine the ultimate accuracy and reliability of the model results. Even under the best circumstances, the model results should be considered as estimates or approximations, since the model itself is an approximation of a real environmental system. This does not detract from the utility of models; it simply emphasizes the use of the model as a tool. It is also a very valuable learning tool - to understand the critical factors that determine the behavior of the simulated system. With this knowledge, the system can be better managed. Most models are more accurate in a relative sense, than in an absolute sense. That is, when models are used to compare alternatives (such as management or control options) the relative differences predicted between alternatives are usually more reliable then the absolute predicted value (or values) for any one alternative. This is a common use of models and one of their major advantages - the ability to project the impact of changed conditions as a basis for evaluating alternative management or regulatory options. When absolute values are needed, such as estimating probable exposure concentrations of a chemical (say, a pesticide) for comparison with drinking water and/or health-effects levels, model results should be supplemented with sensitivity and/or uncertainty analyses in order to analyze the potential "real-world" variability about the model-predicted values. In other words, consideration of the uncertainty of the simulated results is at least as important as are the results themselves in any decision-making process.

In summary, the model user and decision-maker must interact closely throughout the model selection, application, and analysis/interpretation phases, but it is especially critical in the last phase. Only in this way can we insure that reliable information needed for decision-making is produced by the modeling effort.

MODEL TESTING

Model testing was described earlier as encompassing the steps of calibration, validation, and post-audit analyses (when performed). This section explores the model testing process in terms of the critical issues to be considered, the extent to which models have been validated, and the types of generic model validation procedures and guidance needed to bring some order and consistency to this area of environmental analysis.

Model Testing, Validation, and Error Analysis

The process of field validation and testing of models can be viewed as a systematic analysis of errors. In any model calibration/validation effort, the model user is continually faced with the need to analyze and explain differences (i.e., referred to as errors in this discussion) between observed data and model predictions. This requires assessments of the accuracy and validity of observed model input data, parameter values, system representation, and observed output data. Figure 5 schematically compares the model and the natural system with regard to inputs, outputs, and sources of error. Clearly, there are possible errors associated with each of the categories noted above; they can have dramatic effects on the conclusions of the model validation process. Each type of error is described briefly below; in-depth discussions and examples are included in earlier publications (Donigian 1982b, 1983).

Input Errors

Errors in model input often constitute one of the most significant causes of discrepancies between observed data and model predictions. As shown in figure 5, the natural system receives the "true" input (usually as a "forcing function") whereas the model receives the "observed" input as detected by some measurement method or device. Whenever a measurement is made, a possible source of errors is introduced. System inputs usually vary continuously both in space and time, whereas measurements are usually point values, or averages of multiple point values, and for a particular time or accumulated over a time period. Although continuous measurement devices are in common use, errors are still possible, and essentially all models require transformation of a continuous record into discrete time- and space-scales acceptable to the model formulation and structure.

System Representation Errors

System representation errors refer to differences in the processes and the time- and space-scales represented in the model, versus those that determine the response of the natural system. In essence, these errors are the major ones of concern when one asks "How good is the model?".

Whenever comparing model output with observed data in an attempt to evaluate model capabilities, the analyst must have an understanding of the major natural processes, and human impacts, that influence the observed data. Differences between model output and observed data can then be analyzed in light of the limitations of the model algorithm used to represent a particularly critical process, and to insure that all such critical processes are modeled to some appropriate level of detail.

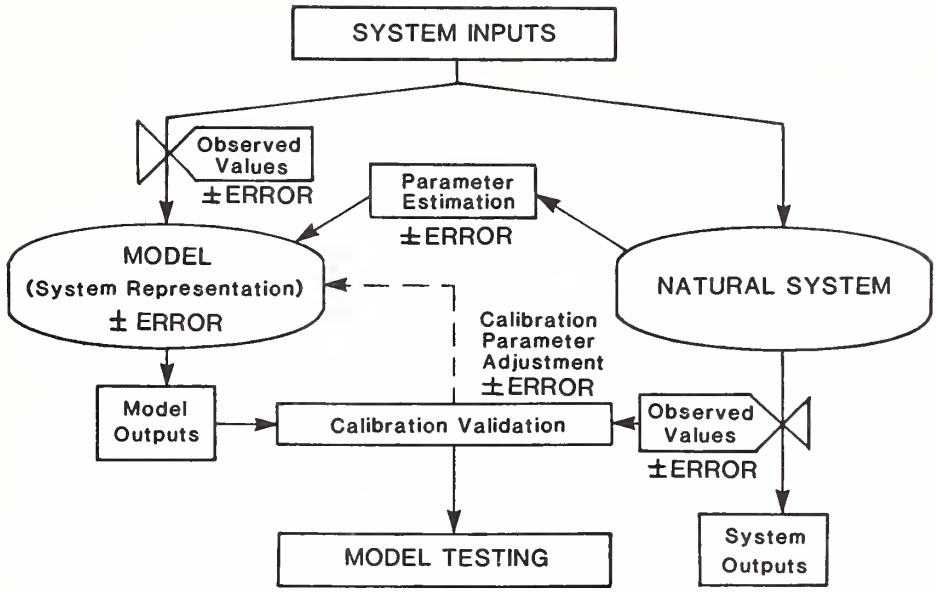


Figure 5.
Model vs. natural systems: inputs, outputs and errors.

Parameter Errors

Parameter errors are caused primarily by the inability to accurately measure and predict many of the parameters that characterize the natural system and, for chemical-fate modeling, the relevant chemical processes under field conditions. Errors are associated with parameter values obtained both from actual measurements and from model calibration. Due to natural variations in topography, soil and aquifer characteristics, crop cover densities, etc., a single parameter value or time-varying parameter function (e.g., crop cover) will always have some associated error. The goal is to obtain measured parameter values that, to the extent possible, represent mean or average conditions over appropriate spatial and temporal scales for the natural system. If the specific parameter values vary significantly, then further segmentation of the system representation may be needed in order to develop domains with relatively uniform characteristics. Such is the case for watershed modeling where different crop types and vastly different soils characteristics are encountered.

Parameters for which measured values are not clearly defined or readily available are often determined through calibration of the model with observed data. Although initial parameter values can always be estimated, calibration is usually recommended to account for local and spatial variations. In many modeling efforts, conscientious model users will often overrun the calibration budget because of the natural tendency to continue to make calibration runs in an effort to minimize discrepancies between simulated and observed values. Parameter errors associated with calibration are often a result of missing and/or erroneous data either as system inputs or outputs, or a result of system representation errors (i.e., inappropriate or inadequate model).

Errors in system output measurements can also produce calibration errors because the model user may be attempting to calibrate against inaccurate or missing data.

Output Errors

Output errors are analogous to input errors; they can lead to biased parameter values or erroneous conclusions on the ability of the model to represent the natural system. As noted

earlier, whenever a measurement is made, the possibility of an error is introduced. For example, published USGS streamflow data, often used for calibration of hydrologic models, can be 5% to over 15% in error; this, in effect, provides a tolerance range within which simulated values can be judged to be representative of the observed data. It can also provide a guide for terminating calibration efforts.

Output 'errors' can be especially insidious since the natural tendency of most model users is to accept the observed data values as the "truth" upon which the adequacy and ability of the model will be judged. Model users should develop a healthy, informed skepticism of the observed data, especially when major, unexplained differences between observed and simulated values exist. It is clearly inappropriate to allocate all differences between predicted and observed values as model errors; measurement errors in field-data collection programs can be substantial and must be considered.

Recent Efforts in Developing Model Testing and Validation Procedures

At the beginning of this decade, a number of symposia and workshops were convened to explore the specific topics of model validation and field testing of exposure assessment models. Donigian (1982a) has attempted to summarize the conclusions and results of a few of these gatherings. Although the workshops had their own specific emphasis and organization format, the common theme was to define the current state-of-the-art of modeling contaminant fate and transport for various media and constituents, to establish the extent of field testing that had been performed, and to communicate future needs in terms of research and data collection for continuing model development and refinement. Model validation was a common topic at these meetings, and the need for commonly- accepted validation measures and procedures was widely recognized.

In March 1982, the EPA Office of Research and Development convened a workshop with the specific objectives to:

(1) assess the state of knowledge on determining the field applicability of laboratory bioassay tests, toxicity studies, microcosm studies, and mathematical chemical exposure models (i.e., the extent to which these methods have been tested/compared with field data)
(2) recommend research objectives and priorities to advance the current level of field testing

Workshop attendees included representatives from EPA research laboratories, universities, and private industry. Working groups were organized with specific responsibility to assess the utility and limits of four different methods (or tools) currently used by EPA and industry for evaluating hazards posed by toxic chemicals:

(1) laboratory toxicity data
(2) microcosm test data
(3) site-specific data
(4) chemical fate and exposure model results

The Exposure Modeling Work Group concluded (U.S. EPA, 1982) that the current extent and/or adequacy of model field testing could only be assessed with respect to the model accuracy required for specific types of regulatory problems. Screening and site-specific assessments for all media were identified as the most likely problems expected under current and future regulatory conditions. Although specific precision and accuracy requirements could not be defined, these two levels of assessment were characterized as follows (U.S. EPA, 1982):

(1) screening - screening level modeling (e.g., chemical pre- manufacturing notice evaluations) is usually accomplished by far-field models and accuracy is expected to be within one order of magnitude

(2) site-specific - more detailed modeling is required for site-specific problems (e.g., waste load allocations, hazardous waste disposal siting) and accuracy is expected to be within a factor of 2-4 for many situations and within a factor of less than 2 for some situations

The current extent of field testing was evaluated for each level since the testing and data requirements for site-specific assessments would be significantly more stringent than for screening purposes.

Table 6 presents the results of the Exposure Modeling Work Group discussions for the screening and site-specific level models, respectively. The assessment is based on a ranking scale between 0 to 100; 0 indicates situations where no testing has been attempted, and 100 identifies areas where extensive testing has been completed with sufficient post-audits to validate the predictive capability of relevant models. The scores can also be interpreted to mean the extent to which additional field testing would improve our understanding of how well the models represent natural systems. The scores do not indicate model accuracy per se; they show the degree to which current field testing has been able to identify or estimate model accuracy.

Despite the recognized need for model validation procedures, concentrated efforts have not been made to develop appropriate model performance criteria and measures. In the past few years, a resurgence of interest in this area has been evident from recent workshops, publications, and technical exchanges on this topic. A few of these indications are discussed below.

The Exposure Assessment Group of the U.S. EPA ORD Office of Health and Environmental Assessment has initiated a study to define and evaluate the level of validation of models currently used by various EPA offices (Versar, Inc. 1987). The overall objectives of the effort are as follows:

(1) To develop working definitions for the common elements of the validation process.
(2) To identify a comprehensive list of the components or steps of the validation process used by modelers of various media.
(3) To review all existing literature on specific models used by EPA, and develop a database for general distribution that documents the level of validation for each model.

These objectives of this study clearly recognize, and seek to remedy, the current deficiencies in model validation terminology, procedures, and evaluation criteria. The study results are intended to complement recently-developed (by EPA) model selection criteria for air, surface water, and groundwater (e.g., EPA (1987a) and to provide broad guidance for the use of models in exposure/risk assessment (Versar Inc. 1987). Since the study is ongoing at this time, its success and the general acceptance of its results and recommended procedures are yet to be determined.

At the same time the EPA Risk Assessment Forum, an upper-level management advisory council, has initiated a Model Validation Project with the mandate to (1) develop an Agency position on model validation in predictive exposure assessments, and (2) identify and propose guidelines for model peer review and/or validation. Some of the issues to be addressed by the Project include:

a. Are Agency-wide guidelines needed for model validation?
b. Can generic procedures be defined and implemented?
c. What are appropriate procedures for model peer review?
d. How should model validation be factored into the risk assessment process?
e. What is the relationship between model validation and uncertainty assessment?

Table 6. Current extent of field testing for screening level and site-specific models.

100 = Model thoroughly field tested
0 = Model not tested

SCREENING LEVEL MODELS	MEDIA						
			Watersheds			Soil Systems	
MODELED PROCESS	Air	Runoff	Streams	Lakes	Estuaries	Unsat. Zone	Sat. Zone
Dispersion-Diffusion	90	--	100	100	100	50	75
Advection	75	100	100	100	100	100	100
Intermedia Transfers							
-sorption/desorp.	75	80	80	80	80	80	80
-volatilization	10	50	100	80	80	70	
-wet/dry deposition							
-to land	50	--	--	--	--	--	--
-to water	90	--	--	--	--	--	--
Sediment/Particulate Transport	75	60	60	90	60	--	--
Transformation Processes							
-chemical	75	--	80	80	80	10	10
-biological	--	--	10	10	10	10	10
-lumped parameter	--	--	--	--	--	60	25

SITE-SPECIFIC MODELS	MEDIA						
			Watersheds			Soil Systems	
MODELED PROCESS	Air	Runoff	Streams	Lakes	Estuaries	Unsat. Zone	Sat. Zone
Dispersion-Diffusion	80	--	90	80	80	10	25
Advection	25	75	100	80	80	15	30
Intermedia Transfers							
-sorption/desorp.	50	40	40	40	20	20	30
-volatilization	10	15	80	20	20	15	15
-wet/dry desposition							
-to land	10	--	--	--	--	--	--
-to water	60	--	--	--	--	--	--
Sediment/Particulate Transport	60	30	20	30	10	--	--
Transformation Processes							
-chemical	45	--	80	80	80	10	10
-biological	--	--	10	10	10	10	10
-lumped parameter	--	--	--	--	--	60	25

Source: U.S. EPA, 1982

In developing a proposed Agency-wide position on model validation, the study will review and attempt to integrate the results of past and ongoing investigations, such as the EAG study (noted above), other efforts discussed below, and the work of other agencies (e.g., Nuclear Regulatory Commission, USGS, FEMA). The study results are intended to include technical guidance for performing model validation for specific types of predictive exposure assessments (L. Mulkey, personal communication 1988).

The International Ground Water Modeling Center is collecting and reviewing information related to the use of performance standards and acceptance criteria for groundwater models used in planning and decision-making. They consider performance criteria in terms of both model validity, related to the scientific correctness and accuracy of the model, and model efficiency, relating to model use of computer resources (e.g., core and mass storage, input/output) and processing time. Acceptance criteria include both management-oriented issues, such as user-friendliness and required user capabilities, model accessibility (e.g., effort, cost, use restrictions), and acceptable temporal and spatial segmentation, and requirements for publication and peer review of complete model documentation (i.e., conceptual and mathematical framework, assumptions and limitations, availability of code testing results) (van der Heijde 1987). Early results of the IGWMC efforts have included publication of proposed benchmark datasets or problems that could be used in validation of groundwater flow and transport models (Huyakorn et al. 1984), and a generalized framework that defines the following three levels of model validation (van der Heijde et al. 1985):

Level I - Code verification on accuracy and operation of computer algorithms, including comparison with analytical solutions and sensitivity analyses.

Level II - Tests of special features and extreme conditions lacking analytical solutions.

Level III - Model validation, including calibration, prediction, and comparisons with field data.

The IGWMC's emphasis on code efficiency, computer operation, and output analysis and interpretation are indicative of the types of issues faced by users in applying complex, multi-dimensional groundwater flow and transport models. This demonstrates the extent to which model validation issues are likely to be both media and application dependent.

In November 1987, the Agricultural Research Institute sponsored a workshop entitled "Research Needs for Unsaturated Zone Transport Modeling of Agricultural Chemicals". The workshop format included presentations on the types of applications of unsaturated zone modeling for management, research, and education, and work group discussions to define research needs and recommended programs for flow processes, soil properties, pesticide processes, nutrient processes, and model validation. The Model Validation Work Group, after considerable discussion, accepted the definitions of 'model', 'model validation', and associated terms established by ASTM (with slight modifications), as discussed earlier. The model validation issues identified by the Work Group included the following:

(1) Acceptance criteria for model validation
(2) Acceptable input parameter estimation procedures
(3) Model enhancements and evolution (user feedback)
(4) Availability of adequate field data sets for validation and comparison

Specific research needs were then identified for each of these issues, the details of which are included in the Model Validation Work Group report (Donigian 1987) included in the workshop proceedings. The discussion on acceptance criteria concentrated on the need for quantitative measures of model performance as a basis for establishing model validation. The formation of a multi-organizational council was recommended, comprised of model developers and users from

government, industry, academia, and consulting firms, to review available information and develop appropriate criteria for different types of model applications and scales.

Procedures for independently estimating chemical and biological rate constants from fundamental soil, environmental, and compound characteristics was identified as a high priority research need. Mechanisms for promoting interactions between model developers, researchers, and model users are needed as part of continuing model evolution and refinement, in addition to providing support for model users. Comprehensive field data sets on which to validate models and provide a basis for model comparisons need to be identified and made available to model developers and users.

Despite a recent emphasis on model testing and validation throughout the literature, unified and accepted procedures and measures for model validation do not yet exist. Although procedures commonly used for model testing are often problem and model specific, three general categories have been identified, (U.S. EPA 1982):

(1) Model parameter estimation by laboratory, microcosm, or pilot plant studies, followed by field application.
(2) "Split-sample" field testing involving calibration and validation on separate data sets, often for different time periods at one site.
(3) Site-to-site extrapolation of model results involving model calibration at one site and subsequent testing against data collected at another site.

These three procedures are often combined in various ways depending on data availability, model structure, and modeling purposes. For example, transport processes may often be calibrated and verified on available data, while the transformation process parameters may be derived from laboratory measurements and applied without calibration.

The greatest need is clearly the use of quantitative measures to describe comparisons of observed and predicted values. Although a rigorous statistical theory for model performance assessments has yet to be developed, a variety of statistical measures has been used in various combinations and the frequency of use has been increasing in recent years. Three general types of comparisons that are often made in model performance testing include (U.S. EPA 1982).

(1) Paired-data performance, involving comparison of predicted and observed values for exact locations in time and space.
(2) Time and space integrated, paired-data performance: spatially and/or temporally integrated data can be compared to analogous model predications, such as daily or monthly averages or totals.
(3) Frequency domain performance involving comparison of cumulative frequency distributions of the observed data and model predictions.

Statistical measures for the paired-data and integrated paired-data performance tests noted above are essentially identical. They include simple statistics (e.g., sums, means, standard deviations, coefficient of variation), error analysis terms (e.g., average error, relative error, standard error of estimate), linear regression analysis, and correlation coefficients. Frequency domain performance has been analyzed with goodness-of-fit tests such as the chi- square, Kolmogorov-Smirnov, and Wilcoxon rank sum tests. The studies by Young and Alward (1983) and Hartigan et al. (1982) demonstrate the use of these tests for pesticide runoff and large-scale river basin modeling efforts, respectively, in conjunction with the paired-data test. James and Burges (1982) discuss the use of the above statistics and some additional tests in both the calibration and validation phases of model testing. They also discuss methods of data analysis for detection of errors; this last topic needs additional research in order to consider uncertainties in the data which provide both the model input and the output to which model predictions are compared.

Green and Stephenson (1986) discussed a number of statistical indices for evaluating the goodness-of-fit (or model performance) of hydrological simulation models to measured hydrograph data from a single storm event. After evaluating 21 different indices, they concluded that "no single statistical goodness-of-fit criterion is sufficient to adequately assess for all purposes the fit between a computed and measured hydrograph." They noted that each index was "weighted in favor of a different hydrograph component (e.g., volumes, peak flow rates)" and because of this the index chosen to assess a model's performance "should depend on the objective of the modeling exercise." These findings suggest that more than one goodness-of-fit criterion be used in a generic evaluation of a model, whereas a single index might be adequate for a model application with a specific objective (e.g., prediction of pesticide mass emissions below crop root zone). Furthermore, in the absence of tests to determine the statistical significance of the differences in their values, these indices can only be used to evaluate the relative performance of models.

The topic of model validation and testing, and many of the concepts discussed above have been included in a number of recent workshops to which the reader is referred for more details (Dickson et al. 1982, Swann and Eschenroeder 1983).

CONSIDERATIONS IN DEVELOPING A MODEL VALIDATION FRAMEWORK

This brief review of the current literature and ongoing efforts related to model validation has shown that the need for a generalized framework for model validation is well recognized in the modeling community, does not exist at the current time, but it is the subject of considerable discussion and ongoing research. Although we would like to be able to satisfy this need by presenting such a framework in this paper, the best we can hope to do is identify some of the key considerations that must be part of a generalized framework. A number of common threads are evident.

Model validation is more a process than a result. According to Hern et al. (1986), "... model development and subsequent validation is an evolutionary process by its very nature." It involves multiple assessments of the model's capabilities to represent observed data under a range of conditions. Indeed, it may be more appropriate to call these individual assessments model testing and evaluation, as steps along the path of model validation. This has been suggested by Green (personal communication 1988), and Versar (1987) who feel that a documented model application constitutes one model evaluation, as part of the validation process.

The term 'validation' itself may be part of the problem; it inherently implies a positive result, i.e., that the model is valid for the conditions simulated. However, negative results showing lack of agreement between a model and observed data are just as valuable, if not more valuable, because they may help to demonstrate the bounds of applicability, or limitations, of the model. This type of information is essential for model users to know when not to use the model results, and/or question its results, and for model developers to identify what parts of the model need to be improved.

Model validation procedures will depend on the media and the type of application. This dependence has been recently recognized by Jones and Rao (1988), Versar (1987), and the ARI Workshop Model Validation Work Group (Donigian 1987). Specific model performance and acceptance criteria, and technical issues, will be different for different media; classifications could include air, land, water, soils (vadose zone), groundwater, or combinations within watershed and soils/groundwater systems. The types of applications may be classified as screening vs. site-specific, or management vs. design vs. regulatory, or some other categorization.

Technical issues that must be considered in model validation procedures, the specific details of which will likely be different for individual media and applications, include the following:

a. Variability in observed data, both input and output, must be recognized and its impact must be considered.

   Both temporal and spatial variability of model input, processes, parameters, and observed data should be evaluated. Spatial variability may be less important for specific types of applications in air and watershed systems, whereas it may be an overriding concern in soil systems. Indeed, it has been suggested that complete validation of unsaturated zone transport models may not be possible partly due to spatial variability issues (Rao and Wagenet 1985). Impacts of variability on the uncertainty of model predictions should be addressed; Jones and Rao (1988) discuss the types of uncertainty involved in the simulation of agricultural chemicals in the unsaturated zone, and Carsel et al (1988) have proposed Monte Carlo-based procedures for quantifying these types of uncertainty. Similar procedures may be available for other media and applications.

b. Parameter estimation procedures must be well-defined and accepted.

   This was a major need identified by the Model Validation Work Group of the ARI Work Shop (Donigian 1987). Jones and Rao (1988) have noted that independently determining degradation rates of agricultural chemicals may be the most difficult problem in separating calibration from validation of unsaturated zone transport models. This is due to the fact that degradation rates are often determined from the same data set used in the calibration. For modeling chemical fate in all media, evaluation of the appropriate transformation rates is a major issue.

c. Benchmark data sets are needed for field testing and model validation.

   The EPA workshop on model field testing in 1982 (U.S. EPA 1982) attempted to identify the types of coordinated data collection programs needed for model validation in air, streams/lakes/estuaries, and runoff/unsaturated/saturated zone media categories. Six years later the ARI workshop re-stated the need to identify benchmark field data sets that could be used for model validation and comparisons. The concept of a 'minimum' field data set was discussed so that guidelines for design of field programs could be established with model validation as the objective. Many participants felt that a number of adequate field data sets existed, but were not known to or accessible by model developers. A type of central clearinghouse to collect and distribute the datasets was suggested to alleviate this problem (Donigian 1987).

d. Performance and acceptance criteria for model validation must be defined.

   This is probably the most commonly recognized and vocalized need in model validation. Appropriate criteria must be both media and application specific, and must include the issues (noted above) related to data variability and uncertainty, parameter estimation, and benchmark or minimum data sets for adequate field testing and model validation.

Model validation procedures will need to include an overall framework that consists of generic issues and requirements for all media/applications, and specific technical procedures that are a function of the specific media and the type of application. Unfortunately, our governmental structures and technical disciplines are not well suited for developing solutions to joint multi-disciplinary, multi-media problems. Our lack of progress in the past in developing a generic framework for model validation procedures may be because we have approached it from the limited perspective of individual media or single applications. Although the media and application specificity is required at the lower-level technical detail, the upper-level guidance and framework must be comprehensive and broad-scale.

Table 7 shows a proposed sample framework for the development of model validation procedures in a multi-level approach. Level I would comprise the generic procedures and guidance (or requirements) applicable to all media and all applications. It would involve clarification of terminology and definitions, minimum standards for model documentation, data comparisons and code verification, peer review requirements, and specification of the issues that must be addressed in establishing performance and acceptance criteria for model validation efforts. These Level I procedures would be developed by the type of multi-organization, multi-discipline group recommended by the ARI Workshop, so that all media and all types of applications would be represented.

Level II procedures would adapt and quantify (where necessary) the Level I procedures to specific media and types of application. The technical issues discussed, such as variability and uncertainty, parameter estimation, benchmark data sets, peer review procedures and performance/acceptance criteria, would be addressed. The Level II procedures would be developed by media-specific experts, including both model users and developers, so that all types of applications would be considered. These procedures would then be reviewed by the Level I group to insure adherence and consistency with the Level I guidance.

REFERENCES

ASTM. 1984. Standard Practice for Evaluating Environmental Fate Models of Chemicals. 1978-84. American Society for Testing and Materials, Philadelphia, PA.

Boutwell, S.H., S.M. Brown, B.R. Roberts, and D.F. Atwood. 1985. Modeling Remedial Actions at Uncontrolled Hazardous Waste Sites. Anderson-Nichols & Co., Inc., Palo Alto, CA., EPA/540/2-85/001. Prepared for Environmental Protection Agency, Hazardous Waste Engineering Research Laboratory, Cincinnati, OH.

Table 7.
Sample framework for model validation guidelines.

Level I	Generic Model Validation Procedures
	Terminology definition
	Recommended/required steps
	Standards for data comparisons
	Documentation standards
	Peer review requirements
	Definition of media and application categories
Level II	Media- and Application- Specific Technical Guidelines
	Benchmark data sets for validation
	Variability/uncertainty assessment procedures
	Acceptable parameter estimation procedures
	Required validation steps/procedures
	Performance/acceptance criteria
	Minimum peer review requirements

Burns, L., R. Lassiter, and D. Cline. 1982. Documentation for the Exposure Assessment Modeling System (EXAMS). EPA 600/3-82-023. U.S. Environmental Protection Agency, Environmental Research Laboratory, Athens, GA.

Carsel, R.F., R.S. Parrish, R.L. Jones, J.L. Hansen, and R.L. Lamb. 1988. Characterizing the Uncertainty of Pesticide Movement in Agricultural Soils. J. Contam. Hydrol. 2:111-124.

Carsel, R.F., C.N. Smith, L.A. Mulkey, J.D. Dean, and P.P. Jowise. 1984. User's Manual for the Pesticide Root Zone Model (PRZM), Release I. EPA-600/3-84-109. U.S. Environmental Protection Agency, Environmental Research Laboratory, Athens, GA.

Dickson, K.L., A.W. Maki, and J. Cairns, Jr. 1982. Modeling the Fate of Chemicals in the Aquatic Environment. Ann Arbor Science, Ann Arbor, MI. 413 p.

Donigian, A.S., Jr. 1982a. Water Quality Modeling in Relation to Watershed Hydrology. In: Modeling Components of Hydrologic Cycle. V. Singh (ed.) Water Resources Publications, Littleton, CO. 343-382 p.

Donigian, A.S. Jr. 1982b. Field Validation and Error Analysis of Chemical Fate Models. In: Modeling the Fate of Chemicals in the Aquatic Environment. Dickson et al. (eds). Ann Arbor Science Publishers; Ann Arbor, MI. 303-323 p.

Donigian, A.S. Jr. 1983. Model Predictions vs. Field Observations: The Model Validation/Testing Process. In: Fate of Chemicals in the Environment. Swann and Eschenroeder (eds). ACS Symposium Series 225. American Chemical Society, Washington, DC. 151-171 p.

Donigian, A.S. Jr. 1987. Report of Model Validation Discussion Group. In: Proceedings of the Workshop on Research Needs for Unsaturated Zone Transport Modeling of Agricultural Chemicals. Annapolis, MD. Nov. 2-4. Agricultural Research Institute, Bethesda, MD.

Donigian, A.S. Jr., and D.C. Beyerlein. 1985. Review and Analysis of Available NPS and Integrated Watershed Models. Prepared for Woodward-Clyde Consultants, Walnut Creek, CA and U.S. Environmental Protection Agency, Washington, DC.

Donigian, A.S. Jr., and B.R. Bicknell. 1986. Environmental Fate of Contaminants from Sludge Disposal Alternatives to Ocean Dumping: Land Application Report. Prepared for U.S. Environmental Protection Agency, Office of Policy Analysis, Washington, DC.

Donigian, A.S. Jr., and S.M. Brown. 1983. A Workshop on Water Modeling Needs and Available Techniques for Synfuels Risk Assessment. Prepared for Georgia Institute of Technology, Atlanta, GA.

Green, I.R.A., and D. Stephenson. 1986. Criteria for Comparison of Single Event Models. Hydrol. Sci. J. 31:395-411.

Hartigan, J.P., Quasebarth, T.F., and E. Southerland. 1982. Use of Continuous Simulation Model Calibration Techniques to Develop Nonpoint Pollution Loading Factors. Proc. Stormwater and W. Q. Model Users Group Meeting.

Hern, S.C., and S.M. Melancon (ed.) 1986. Vadose Zone Modeling of Organic Pollutants. Lewis Publishers, Inc. Chelsea, MI.

Hern, S.C., S.M. Melancon, and J.E. Pollard. 1986. Generic Steps on the Field Validation of Vadose Zone Fate and Transport Models. In: Vadose Zone Modeling of Organic Pollutants, S.C. Hern and S.M. Melancon (eds.), Lewis Publishers, Inc. Chelsea, MI.

Huyakorn, P.S., A.G. Kretschek, R.W. Broome, J.W. Mercer, and B.H. Lester. 1984. Testing and Validation of Models for Simulating Solute Transport in Ground Water: Development, Evaluation, and Comparison of Benchmark Techniques. International Ground Water Modeling Center, GWMI 84-13, Holcomb Research Institute, Butler University, Indianapolis, IN. 419 p. November.

HydroQual, Inc. 1985. Summary of Available Models for Receiving Water Impacts from Nonpoint Sources. Prepared for Woodward-Clyde Consultants, Walnut Creek, CA. and U.S. Environmental Protection Agency, Washington, DC.

James, L.D., and S.J. Burges. 1982. Selection, Calibration, and Testing of Hydrologic Models, in Hydrologic Modeling of Small Watersheds. ASAE Monogr. 5. C.T. Haan, H.P. Johnson, D.L. Brakensiek, Eds. ASAF, St. Joseph, MI. Chap. II.

Jones, R.L., and P.S.C. Rao. 1988. Reflections on Validation and Applications of Unsaturated Zone Models. Prepared for the Workshop on the Validation of Flow and Transport Models, Ruidoso, New Mexico, May 23-26.

Kincaid, C.T., J.R. Morrey, S.B. Yabusaki, A.R. Felmy, and J.E. Rogers. 1984. Geohydrochemical Models for Solute Migration, Volume 2, Preliminary Evaluation of Selected Computer Codes for Modeling Aqueous Solutions and Solute Migration in Soils and Geologic Media, EA-3417, Volume 2, Electric Power Research Institute, Palo Alto, CA.

Lorber, M.N., and L.A. Mulkey. 1982. An Evaluation of Three Pesticide Runoff Loading Models. J. Environ. Qual. 11(3):519-528.

Lumb, A.M., and J.L. Kittle, Jr. 1986. ANNIE - An Interactive Processor for Hydrologic Modeling. Water Resources Investigation Report 83-xxxx (DRAFT). USGS, Surface Water Branch, Reston, VA.

Rao, P.C.S., and R.J. Wagenet. 1985. Spatial Variation of Pesticides in Field Soils: Methods for Data Analysis and Consequences. Weed Sci. 33(2):18-24.

Swann, R.L., and A. Eschenroeder. 1983. Fate of Chemicals in the Environment. ACS Symposium Series 225, ACS, Washington, DC. 320 p.

U.S. Environmental Protection Agency. 1987a. Selection Criteria for Mathematical Models Used in Exposure Assessments: Surface Water Models. Exposure Assessment Group, Office of Health and Environmental Assessment, ORD. EPA-600/8-87-042.

U.S. Environmental Protection Agency. 1987b. Selection Criteria for Mathematical Models Used in Exposure Assessments: Ground Water Models. Exposure Assessment Group, Office of Health and Environmental Assessment, ORD.

U.S. Environmental Protection Agency. In Preparation. Selection Criteria for Mathematical Models Used in Exposure Assessments: Air Quality Models. Exposure Assessment Group, Office of Health and Environmental Assessment, ORD.

U.S. Environmental Protection Agency. 1982. Exposure Modeling Committee Report. Testing for the Field Applicability of Chemical Exposure Models. Proc. Workshop on Field Applicability Testing. U.S. EPA, Athens, GA.

van der Heijde, P.K.M. 1987. Performance Standards and Acceptance Criteria in Ground Water Modeling. Ground Water Modeling Newsletter, Volume VI(2). 1 p.

van der Heijde, P.K.M., P.S. Huyakorn, and J.W. Mercer. 1985. Testing and Validation of Ground Water Models. International Ground Water Modeling Center. GWMI 85-07, IGWMC Ground Water Modeling Notes, Holcomb Research Institute, Indianapolis, IN. 17 p. August.

Versar, Inc. 1987. Current and Suggested Practices in the Validation of Exposure Assessment Models. DRAFT REPORT. EPA Contract No. 69-02-4254, Work Assignment No. 55. Prepared for U.S. EPA Office of Health and Environmental Assessment, Exposure Assessment Group, Washington, DC.

Yeh, G.T. 1981. AT123D: Analytical Transient One-, Two-, and Three- Dimensional Simulation of Waste Transport in an Aquifer System. ORNL-5602. Oak Ridge National Laboratory, Oak Ridge, TN.

Young, G.K., and C.L. Alward. 1983. Calibration and Testing of Nutrient and Pesticide Transport Models, in Agricultural Management and Water Quality. Iowa State University Press, Ames, IA. 267 p.

DISCUSSION OF PAPERS PRESENTED IN TECHNICAL SESSION 6,PART 2: PREDICTION/COMPARISON - SIMPLE/COMPLEX MODELS

David Bowles[1], Presiding
Chris Duffy[2], Recorder

PAPERS DISCUSSED

Complexity and Uncertainty in Predictive Models by K.J. Beven and A.J. Jakeman

Selection, Application, and Validation of Environmental Models by A.S. Donigian, Jr., and P.S.C. Rao

SPECIFIC QUESTIONS AND COMMENTS

Questions: (K. Baun, Wisconsin Department of Natural Resources) Dr. Donigian, I have a distributed parameter model which has been through a validation testing and which seems to look okay. Based on the model, we made field specific decisions about who gets funding and who doesn't. Are practices going to be applied or aren't they? People come back to me and say: "Granted, you've done your testing at a set water level, but how confident are you on sight-specific or field-specific output from the model?" Frankly, I don't know. Do you have a response on a calibration procedure for a distributed model like this? How can you ever tell whether the specifics are equal to the sum of the parts?

Response: (A. Donigian, Aqua Terra Consultants, Mountain View, California) I guess I'd say it's a matter of scale. If you're making decisions at the field level, then your model needs to be sensitive to that. You probably need to get another segment in your model so you can come up with decisions at the field scale, and basically tell the farmer: "Yes, we may not be able to identify areas on a field scale, but at least we can at the general region level," so he knows he's within that region. If you're applying the model at a watershed scale with very gross segments (larger segments) you may not be able to identify specific areas within a segment as being contributors or how management practices should be applied at that level. So, it's a matter of modeling scale I believe. At least in my opinion.

Comment: (K. Baun) The scale is quite detailed so field data can be used at a watershed level. But I really don't know how confident I can be at the output for the individual fields. Collectively, it calibrates well; but I can't place any confidence in it for each field.

Question: (A. Donigian) And you have known data at the field scale level?

Response: (K. Baun) No, not really. We're looking at delivery from field to field through ephemeral channels and into the receiving water. How much of it is coming from each particular field?

[1]David Bowles, Professor and Associate Director, Water Research Lab, Utah State University, Logan, Utah.
[2]Chris Duffy, Department of Agronomy, Cornell University, Ithaca, New York.

Response: (A. Donigian) That's a problem of the scale of application. You really need to have data at the field scale to convince someone that it's performing at the field scale as well as what it's doing at the watershed scale.

Question: (D. Gustafson, Monsanto, St. Louis, Missouri) I have a question regarding the ASTM definition of validation. All the definition said was that there is a comparison of the model predictions with the data, it said nothing about how that comparison should be evaluated. Is the EPA committee that you're on going to expand on this? For instance, it seems to me that at least we should request that the residuals average to zero, much as Tony showed in his talk. Are any sort of criteria going to be established as part of the validation?

Response: (A. Donigian) That's a key aspect of it. It was brought up in the workshop and in a number of articles in the literature. We have to have performance criteria and acceptance criteria for each media and each type of application: what are these criteria? Is a correlation coefficient of .8 good enough? We'll probably be relying, as I see it, on experts within the media, within that particular technical area, to come up with these types of guidelines. They will then be reviewed, hopefully, by a wider audience. How do we know the model is valid, unless we have something to compare it to, or some standard. On another point: though it's not clear to me that one measure will always or can always be used for all models, or even for a specific media. You may need two or three different measures. There have been a number of noted researchers who really question whether some models can be validated at all for a general application. So, some of these issues will have to be considered. And I suspect it won't be a correlation coefficient or a standard error of the estimate (one value), there will probably be a range of types of tests that need to be made, and hopefully they will provide a range of acceptance within these various tests.

Question: (D. Gustafson) Should we be using models that can't be validated?

Question: (A. Donigian) What are the options?

Response: (D. Gustafson) Throw away the model.

Question: (A. Donigian) And what do you do on the decision?

Question: (D. Gustafson) Use models which can be validated. Obviously any model predicts something. Why can't we measure that something and compare it to what the model gives you?

Response: (A. Donigian) Yes, but you also have certain bounds of error: how close do you have to be? These are the questions of model validation. How close do you have to be? If a model can't be validated and shows to be invalid I'd say, don't use it. But hopefully there is an alternative or fall back procedure to use, another model that may be, in your judgement, more valid.

Comment: (Audience) In response to both speakers, I'd like to say that we always have some particular attribute, or piece of data, that is a part of the system. It is some behavior criteria associated with the system, and in this sense can aid in looking at the level of ambiguity in a given model in reproducing that behavior. Clearly, the more behavior you've got that's measured with respect to your system, the more you're going to be able to reduce the ambiguity. But still one can always look at it and say whether or not that ambiguity is acceptable, given the behavior criterion you're establishing; whether it's a certain variance in model predictions or whatever.

Question: (S. Glasser, Forest Service, Atlanta, Georgia) I'm not a modeler or a model user, but I'm just wondering if there's any chance for this modeling community to be able to evaluate and then possibly warranty, so to speak, models for different classes of model users. Analogous to the automobile industry; if a model is on the market, different classes of users will come in. Not

everyone is in the market for a Ferrari or would even know how to drive a Ferrari at 150 miles an hour. But, they sure want a Chevrolet, and they know how to drive it 70, 80 miles an hour. I would anticipate that perhaps there are different classes of models that fit different kinds of users and in this mix and match world that we live in, it looks like half of the problem is knowing how to mix and match the models with the users. I, for one, would encourage Dr. Donigian to try and devise a scheme, perhaps in terms of classes of difficulty of model use, to compare the models with classes of users in a generic sense. And in those cases where models are not fully verified and calibrated, that there be some kind of system set up within the modeling community that would ensure that users don't keep on using models that are known to be inaccurate. Would either of you like to comment on that?

Response: (A. Donigian) To set up a scheme such as you are suggesting is almost an impossibility because of a lack of knowledge about user qualifications. A large number of models are put out for general distribution, and they're supported by EPA and USGS and distributed to anyone who wants them. If we can impose a test on the model user, so he has to meet certain capabilities and requirements, before he uses the model, then we might be willing to warrant our models. But we're not going to warrant it, not knowing who the user is and how he's going to use it. An axiom in the modeling community is that once you distribute a model, the first use is probably going to be a misuse, a disregard for the assumptions and limitations of the model that hopefully were clearly stated in the documentation. Another aspect is user support for the models. This has come up as a key issue in a recent workshop and a number of other meetings; and it is an ongoing effort. It was brought out by Ken Thornton yesterday. It's not really a model validation issue, it's basically someone to go to when you have problems. Hopefully all models will have this type of "home", a place to go to so you can ask questions - get the latest version - that's why we have versions 5.1, 6.2, etc. for different models; they're continually evolving. The development of models is an evolutionary process. It is just as important to know what a model can't do as what it can do because it helps the developer to know that these are limitations and need to be improved. These limitations help the user to know what situations he shouldn't apply the model to. So I think a model home or user support system is critical in that type of modeling effort.

Comment: (P. Engesgaard, Danish National Agency of Environmental Protection) I'm getting somewhat concerned because Dr. Donigian presented us with a model for the selection of models. I think it looks extremely complex, even though some people have advocated use of simple models. It also seems to me that it's not a validated model at all, and that is probably why it's presented here; to get some kind of response as to whether this model is feasible for selection of models. Well, I would be very uneasy about using this system if I were presented with this model to select a model. I think things could be simplified a lot more; I believe you'll get confused by using this kind of model. The bottom line is that we do not dare use a model that we do not trust. It appears to me that you're doing a lot on "How do these things work? How well are they documented?" and all that kind of stuff. Before, when somebody asked you, "Well what if it's not validated? What else should we do, do we have to use it anyway?" Then you will say "Yes because there will be a little bit of data that we can always use to validate." But the problem is that what we usually use these models for, in the decision situation, is checking pretty drastic alternatives that are not in the existing data. This is particularly true when you get into biological and chemical systems; you can be way off with something that's calibrated on an existing system. It works fine, because all of you are water people, you're engineers, hydrologists, the models work fine because you know equations of flow. But we don't know the equations. We don't know these things, and we can get way off. And that is a key issue about model selection. Can we trust the model for drastic future changes that we are facing?

Response: (A. Donigian) I disagree. I think there are cases where we can use models. And I think there are many types of applications where we can use them. I'm not saying modeling should be used for everything. It's not in my background to say that all the biological algorithms are 100% correct. But there are cases where models can be used and there are cases where

models cannot be used. In the model selection procedures, I asked for the specifications of your model, what your model has to do? If your model has to represent biological processes, then that will feed into the model selection process and there may be only a few models that will even attempt to simulate such processes, or there may be no models that will even attempt to do it, in which case, you're right, don't use the model; rely on the experts. But I'm saying that there is a role for models in the decision making process, in providing input. And all we're presenting is a kind of step-by-step procedure for coming up with the specifications of what this model has to do to be used effectively.

Comment: (T. Jakeman, Australian National University, Canberra, Australia) I'd also like to say that there's no substitute for analysis; a biologist must have some feelings as to what some of the processes must be and he should try and look at how variability across the spectrum of possible processes will effect the model prediction and he should try and quantify that. If the predictions are between very extreme boundaries, we should know that. By doing that sort of analysis we are going to pinpoint areas of future research. It's the only way we're going to go forward; we shouldn't just throw our hands up in horror and say we can't do it.

Question: (D. Jackson, Susquehanna River Basin Commission, Harrisburg, Pennsylvania) Someone has emphasized the fact that there are errors in measurements which are being used to calibrate these models. We have reservations about sampling procedures that are used to collect field data which may be used for model verification. A key question here seems to be, what is the error in the field measured values? This is both a matter of sampling procedure in the field and in laboratory procedures for determining concentrations. I'd like to ask, what do we know about the accuracy of field sample values, particularly for agricultural chemicals, such as nutrients?

Response: (A. Donigian) You're probably asking the wrong person. I know there are people out there who can probably answer this a lot better than I can. From the earlier presentations, for example in the unsaturated zone, you saw error bounds or confidence limits about each data point. Looking at the issues of spatial variability at a field level, concentrations can vary by orders of magnitude, for a variety of reasons. We need to analyze this data and come up with average field values if we're not specifically trying to represent the spatial variability issue. The other types of errors, in terms of the Geological Survey, extreme flow measurements are usually fairly good; they have 5% - 10% variation on their mean daily and monthly values. These types of variances need to be brought into the modeling process when comparing a model with observed data, to know what the variation is. I'm sure there are probably other people who can better respond to the kinds of error variation you see in chemical data, chemical concentrations? How representative is a particular grab sample or an integrated sample across the stream in representing the model output? The model may be giving you an average value for a stream segment which may be miles long. Whereas, you're measuring at a point in the stream. These issues have to come up in the comparison. I'd welcome any comments from others in terms of variabilities and errors in the measurements.

ANALYZING STATISTICAL PROPERTIES OF NONPOINT-SOURCE WATER-QUALITY VARIABLES

Jose D. Salas[1] and Jim C. Loftis[2]

ABSTRACT

Nonpoint-source water-quality variables are described statistically, and procedures and models for analyzing and synthesizing such variables are discussed. In addition, methods for detecting changes in water quality parameters and corresponding design criteria are reviewed.

INTRODUCTION

Effects of agricultural nonpoint sources on water quality may be described by a number of variables reflecting the use of pesticides, fertilizers and other agricultural practices. These practices will eventually be reflected in the quality of water in rivers, canals, ponds, lakes, reservoirs, soil water stored in the unsaturated zone, and groundwater. To quantify the effects of agricultural practices on water quality one normally measures certain properties or parameters over time and space. This immediately poses the problem of determining how many samples are needed to estimate, for instance, loads of a certain parameter in a given time interval or to detect changes over time of the same parameter. This in turn requires some knowledge of the variability of the underlying variable.

In reality, this analysis may be approached in a broader context by considering the physical bio-geo-chemical processes underlying the water quality constituents as they travel through the hydrologic environment. One may look at the agricultural-hydrological environment in a systems context and analyze the uncertainty of the output as a function of the corresponding input uncertainties and the transfer function model which translates such inputs into outputs. Excellent reviews along these lines have been made by Beck (1987) and Plate and Duckstein (1988). In this paper we do not deal with the total system; rather we focus on analyzing the output.

We discuss three specific topics in this paper: estimation of water quality loads, detection of changes in water quality variables, and sampling of these variables. Inherent in these topics are the statistical characteristics shown by the data and the statistical/stochastic models that may be useful for describing such characteristics. These last two topics are discussed first in the paper since they constitute the basis of the first three.

STATISTICAL PROPERTIES OF WATER QUALITY DATA

Water quality processes in general and nonpoint-source water quality processes in particular evolve on a continuous time scale. However, most of the data readily available for analysis and modeling of such processes have been obtained by sampling the continuous process at discrete points in time. For instance, a daily series of a given water quality variable may be derived by sampling the stream or aquifer once daily. Then, a weekly or monthly series may be obtained

[1]Jose D. Salas, Professor, Hydrology and Water Resources Program, Department of Civil Engineering, Colorado State University, Fort Collins, CO.

[2]Jim C. Loftis, Associate Professor, Department of Agricultural and Chemical Engineering, Colorado State University, Fort Collins, CO.

by aggregating or averaging the daily series over the specified time interval. In any case, series of water quality processes are usually defined at daily, weekly, monthly and annual time intervals. In addition, the term "seasonal" generally refers to monthly time intervals. In this paper, however, it will be used rather loosely, i.e., simply meaning that the time interval is a fraction of the year.

The plot of a water quality series versus time gives a good indication of some of the statistical characteristics. For instance, figure 1 shows the monthly time series of water specific conductivity, at the Yampa River, Colorado. The series shows a typical pattern in which conductivity is low during some months and high during others. This characteristic behavior indicates that the monthly mean varies periodically throughout the year. The stochastic component of the series is also noted beyond the cycle in the mean, although it is difficult to identify any distinct seasonal properties without further analysis. In addition, the monthly series has several gaps, which are characteristic of most water quality series. It also shows a few observations at the end of the record which significantly depart from the rest of the record. This may indicate that a significant change has occurred, or that some outlying observations are present perhaps due to error in measurement. Figure 2 gives another example of time series of monthly water conductivity at Twin Lakes, Colorado (Twin 2). As in the previous example, the data have several gaps and they show either a gradual decrease or a sudden change in the mean of the series. In this example it is difficult to see the annual cycle in the mean without further analysis.

In general, the time series structure of nonpoint-source water-quality variables may be considered as a combination of three components. These are the tendency or trend component, cyclic or periodic component, and the stochastic or random component. The first two may be modeled deterministically, i.e., their values may be assumed known or their future values may be somewhat predictable. The stochastic component is modeled nondeterministically. A heuristic functional representation of these components is

$$X_t = f(T_t, P_t, Z_t) \qquad [1]$$

where X_t = observed time series variable at time t, T_t = trend component,
P_t = within-year periodicity or cyclic component, and
Z_t = stochastic or random component.

Modeling of these components is discussed further in the following sections.

Two types of trend are commonly detected in nonpoint-source water-quality processes: (1) fairly smooth and gradual trends (linear or nonlinear) and (2) abrupt jumps or slippages. Linear or nonlinear trends may be due to long-term changes in agricultural practices and agricultural land development. Jumps or slippages may be created by agricultural activities such as sudden changes in the use of certain types of pesticides, or by large disruptions in nature such as extreme weather or catastrophic events. It is also possible to have a mixture of slippages and jumps with intervening trends (see figures 1 and 2).

Periodicity or cyclicity, refers to within-year (seasonal) variations of the statistical parameters of the underlying variable. Two main factors produce within-year variations in nonpoint-source water-quality variables: (1) the annual hydro-meteorological variables and (2) the usual annual cycle of agricultural activities (for instance, the use of pesticides and fertilizers at certain times through the irrigation season). Hence, time series of water quality variables may be considered to be seasonal in one or more characteristics of the observed data. For example, the mean, variance, skewness and serial correlation.

The stochasticity of observed nonpoint-source water-quality processes is due to inherent randomness and space-time dependence of natural hydrologic processes as well as measurement error.

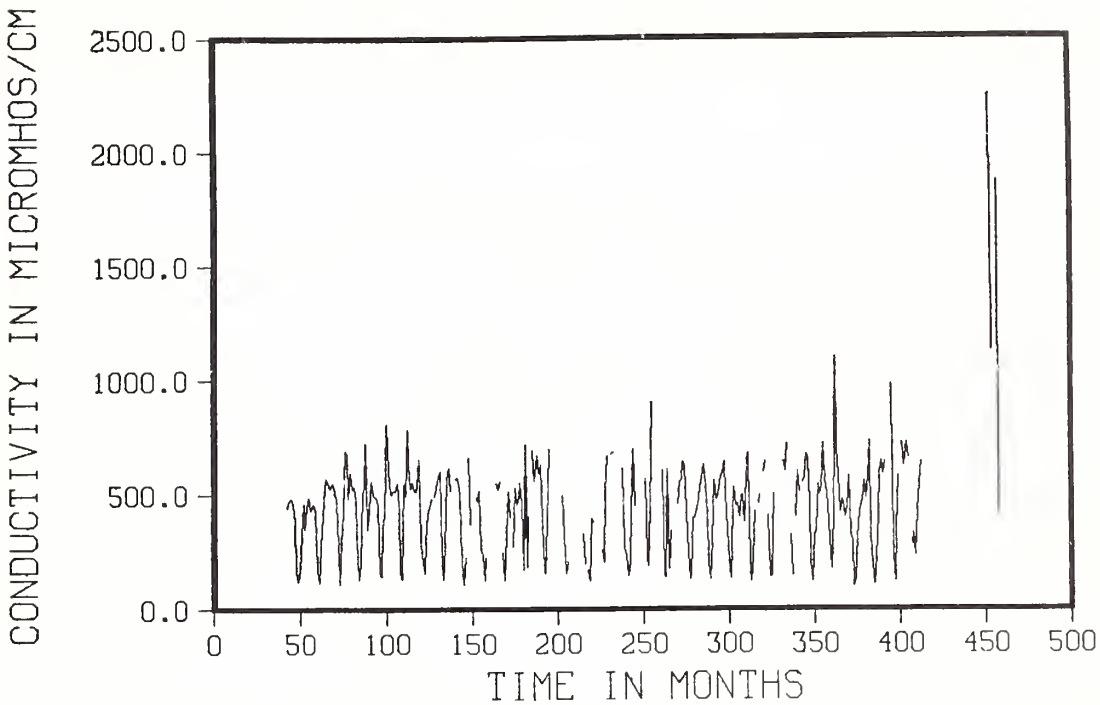


Figure 1.
Example of time series of monthly water conductivity, Yampa River, Co.

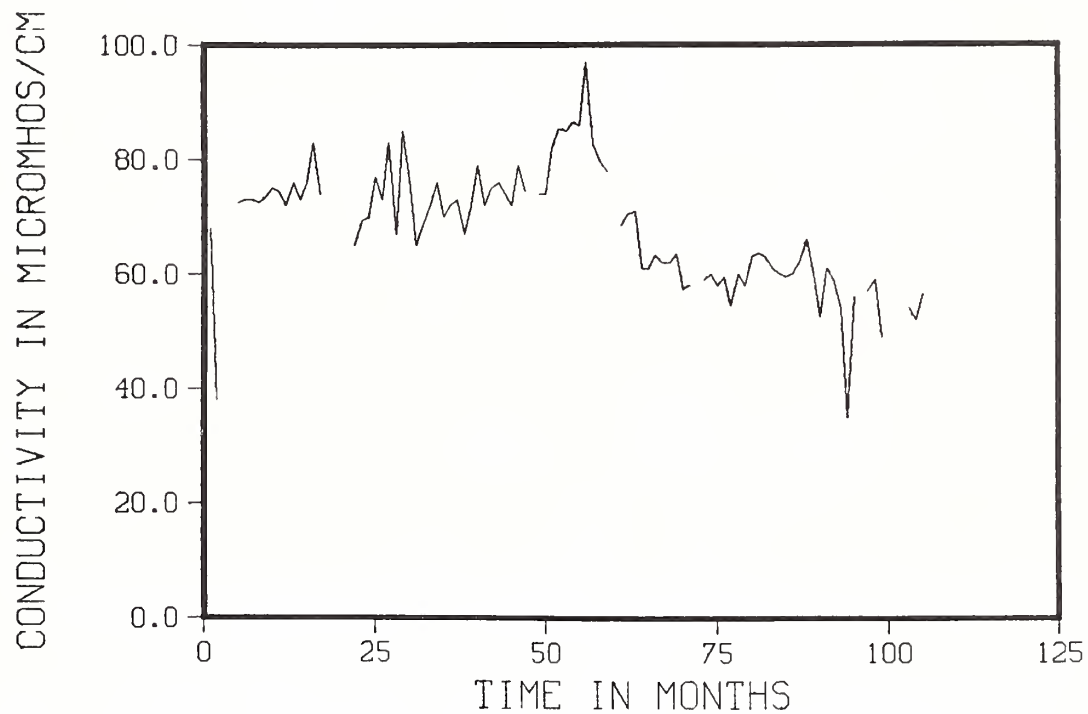


Figure 2.
Example of time series of monthly water conductivity, Twin Lakes, CO (Twin 2).

We next expand upon two of the main statistical properties of non-point source water quality processes: seasonality and dependency structure. Tendency, or trends, is included in a separate section under the general heading of detection of changes.

Seasonality

We will assume that the time series process is trend free or that any significant trend has been previously removed from the original data. It is further assumed that the data consist mainly of periodic and stochastic components. Part of the periodicity can be removed from the original process by simple procedures, i.e., seasonal standardization, while other periodic properties remain and must be incorporated in modeling the stochastic component. In general, nonpoint-source water-quality variables sampled more frequently than once a year may show periodicity in the mean, variance, covariance, skewness and kurtosis. The periodicity in the mean may be easily observed in the plot of the underlying water quality time series. Often the periodicity in the variance (standard deviation) may be also observed from the plot of the time series. However, the periodicity in the covariance is not so easy to observe and usually requires further mathematical analysis. Likewise, periodicity in higher order moments may not be observable from time series data unless further statistical analysis is performed. Actually, not all the characteristics shown by the data may have practical significance. The mean, variance and covariance are the key characteristics considered; kurtosis is rarely included in any analysis.

Let us consider a seasonal water quality time series represented by $X_{\nu,\tau}$, where $\nu=1,2,...,N$ denotes the year; $\tau=1,2,...,\omega$ denotes the season, and ω and N denote the number of seasons in a year and the number of years of record available, respectively. Then, $X_{\nu,\tau}$ represents the observed water quality variable, for example conductivity, during month τ of year ν. The sample seasonal mean $\bar{X}_\tau$, variance S_τ^2, skewness coefficient G_τ and kurtosis coefficient K_τ of the series $X_{\nu,\tau}$, may be determined respectively by (Salas et al. 1980):

$$\bar{X}_\tau = \frac{1}{N} \sum_{\nu=1}^{N} X_{\nu,5vsb0T\tau;\tau} = 1,2,...,\omega \qquad [2]$$

$$S_\tau^2 = \frac{1}{N-1} \sum_{\nu=1}^{N} (X_{\nu,\tau} - \bar{X}_\tau)^2; \qquad \tau = 1,2,...,\omega \qquad [3]$$

$$G_\tau = \frac{N \sum_{\nu=1}^{N} (X_{\nu,\tau} - \bar{X}_\tau)^3}{(N-1)(N-2)\, S_\tau^3}; \qquad \tau = 1,2,...,\omega \qquad [4]$$

and

$$K_\tau = \frac{N^2 \sum_{\nu=1}^{N} (X_{\nu,\tau} - \bar{X}_\tau)^4}{(N-1)(N-2)(N-3)\, S_\tau^4}; \qquad \tau = 1,2,...,\omega \qquad [5]$$

Some examples of periodic water quality statistical properties follow. Figure 3 gives the mean and standard deviation of monthly EC of the Yampa River showing well defined seasonality. Both parameters are periodic, but their behaviors are opposite in that the standard deviations are smaller during the months of large means and vice versa. In other cases (such as streamflow or precipitation) one generally observes smaller standard deviations during months of smaller means. Figure 4 gives the monthly skewness and kurtosis coefficients. Both of these coefficients are also seasonal. In general one may expect that as the order of the statistical moments increases, the periodic pattern becomes less smooth or more "noisy." Likewise, as the sampling frequency of the given variable increases the periodic function becomes more noisy (Yevjevich 1972a). This is why

some kind of smoothing function (for instance Fourier series analysis) is often used in modeling the periodic components of water quality time series.

In addition to the plots of seasonal statistical characteristics as an aid for observing periodicity, the techniques of autocorrelation and spectral analysis may be used for identifying cycles or periodic components. For instance, redefining the original variable $X_{\nu,\tau}$ as simply X_t -- where t denotes the time interval as months, weeks or days -- the lag-k autocorrelation function, r_k, is given by (Box and Jenkins 1970):

$$r_k = \frac{\sum_{t=1}^{N-k} (X_t - \bar{X})\,(X_{t+k} - \bar{X})}{\sum_{t=1}^{N} (X_t - \bar{X})^2} \qquad [6]$$

where $k < N$ is the time lag and $\bar{X}$ is the sample mean. The plot of r_k versus k is called a correlogram. For instance, the correlogram of monthly series of conductivity at Twin Lakes, Colorado (Twin 4) is shown in figure 5. It shows a distinct periodicity of 12 months due to the effect of the annual cycle.

Likewise, spectral analysis has been widely used for detecting cycles in a number of geophysical time series. Wastler (1963) was one of the first to spectrally analyze water quality variables. The spectrum of a given series may be determined by transforming the autocorrelation function r_k as (Jenkins and Watts 1969):

$$g(f_j) = 2\left[1 + 2\sum_{k=1}^{m} D_k\, r_k \cos(2\pi\, f_j\, k)\right] \qquad [7]$$

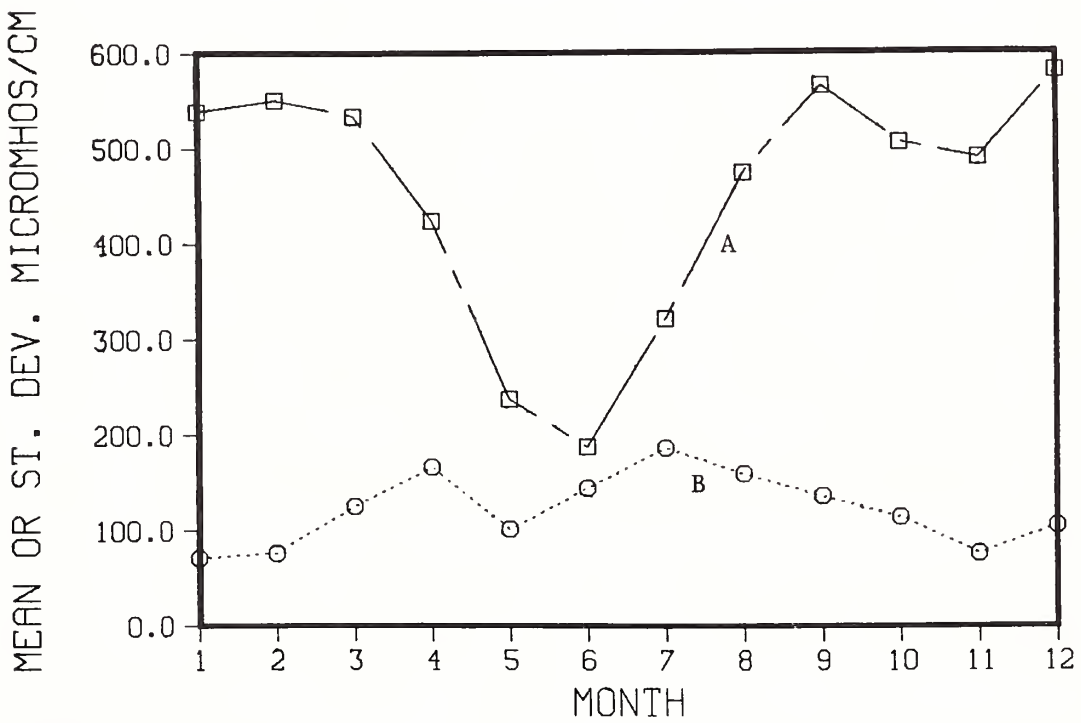


Figure 3.
(A) Mean and (B) standard deviation of monthly water conductivity, Yampa River, CO.

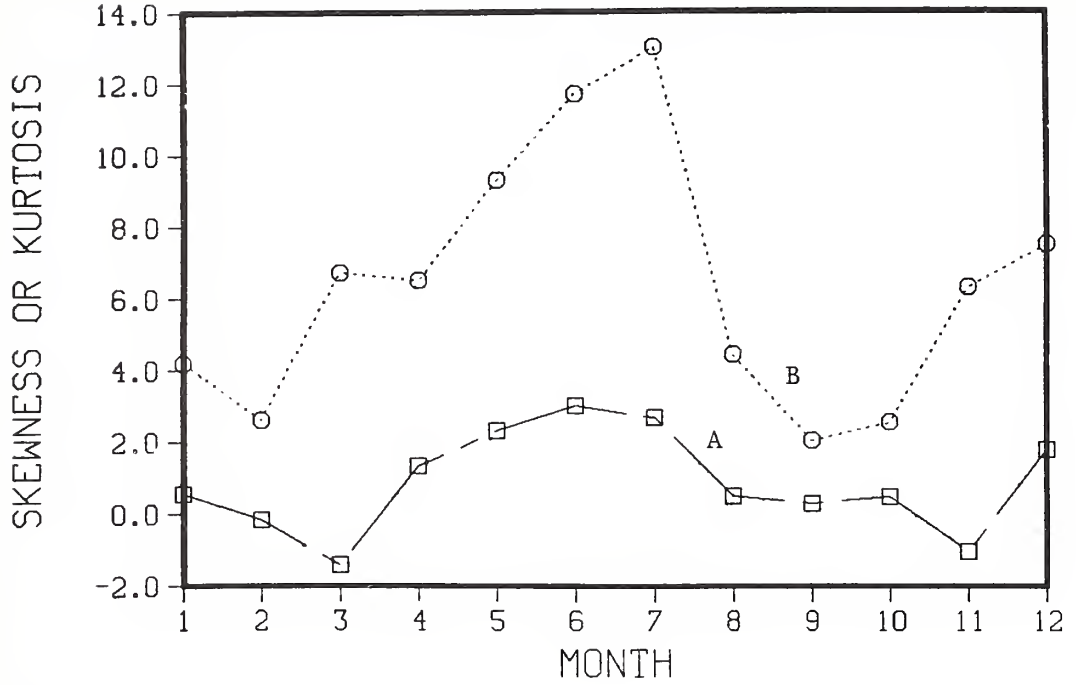


Figure 4.
(A) Skewness and (B) kurtosis coefficients of monthly water conductivity, Yampa River, CO.

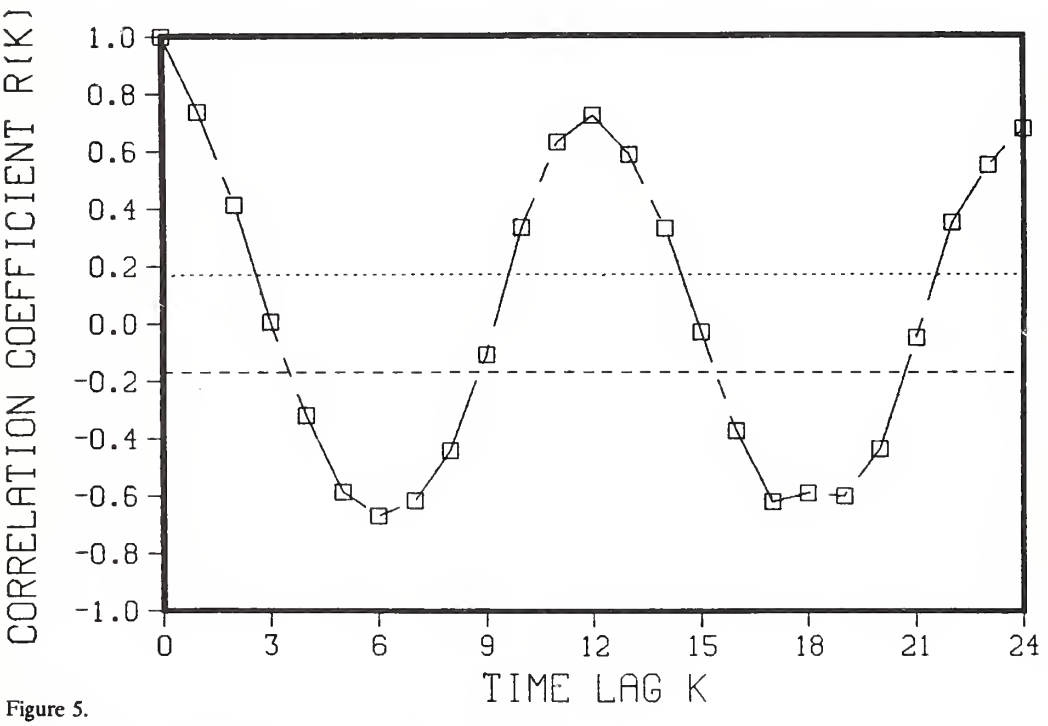


Figure 5.
Correlogram r_k of equation 6 for monthly water conductivity, Twin Lakes, Co (Twin 4).

in which $g(f_j)$ is the spectral density function, $f_j=j/2m$, $j=1,...,m$ is the frequency, m is the maximum number of lags used in the correlogram and D_k is a smoothing function (see for

instance Parzen 1967). As an example, figure 6 gives the spectrum of weekly dissolved oxygen (DO) series for Vermillion River, Louisiana showing a main spike at the frequency of 0.0192 cy/wk, which corresponds to 1 cy/yr (a period of one year), and a smaller but significant spike at the frequency of about 0.0385 cy/wk, which corresponds to 2 cy/yr. Spikes may be tested for significance using Fishers Statistic (Brockwell and Davis 1986).

The spectrum as shown in figure 6 may serve for two (closely related) purposes. One is in representing the time series X_t in the form of a deterministic component D_t plus a stochastic component ϵ_t. The spectrum may suggest that D_t be written in trigonometric form, with a certain number of frequencies (two frequencies in the example above). Lee (1972) used this approach to develop models for predicting daily water temperature, specific conductance, and flow of the Ontario River, Canada. The second purpose is in representing the statistical characteristics such as the mean (or any other characteristic as above defined for that matter) parametrically. The significant spikes in the spectrum may indicate the number of harmonics that may be needed in representing the given seasonal statistical characteristic (usually the mean) in trigonometric form.

One of the simplest methods of removing seasonality is by seasonal standardization, which involves estimating the periodic means and variances of the original series (trend-free series). The series $X_{\nu,\tau}$ defined for season τ of year ν is transformed by

$$Z_{\nu,\tau} = \frac{X_{\nu,\tau} - \bar{X}_\tau}{S_\tau} \qquad [8]$$

where $\bar{X}_\tau$ and S_τ are the seasonal mean and standard deviation given by equations 2 and 3, respectively. This transformation yields a series $Z_{\nu,\tau}$ which may be assumed to be stationary in the mean and variance. This is called the nonparametric method of cyclic standardization wherein raw estimates of the means and standard deviations are used (Yevjevich 1972b). However, the so-called parametric method of seasonal standardization may likewise be used. This involves

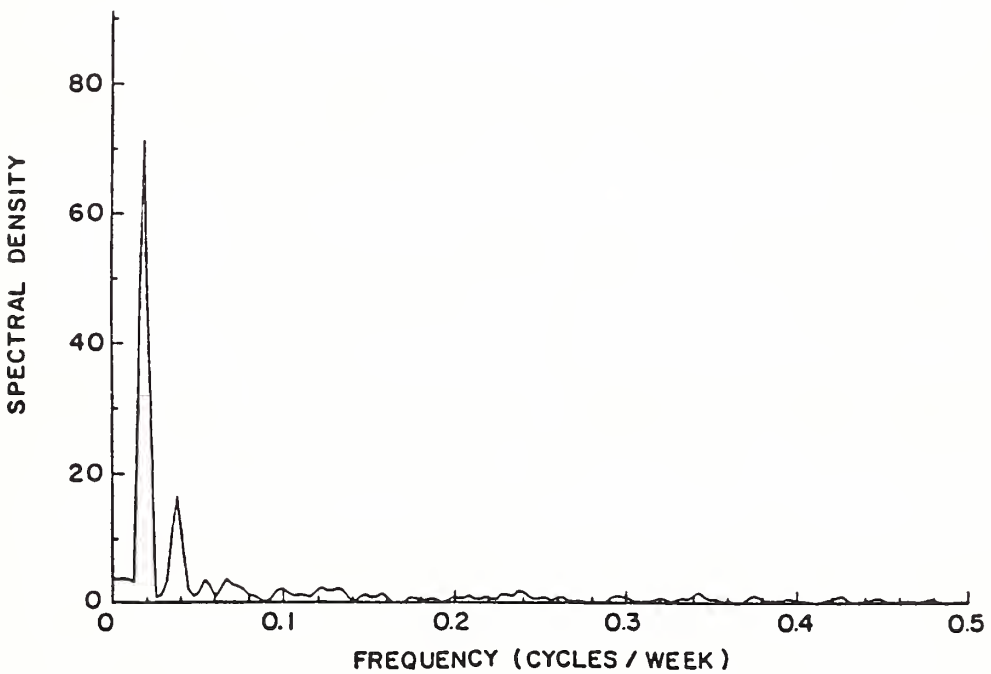


Figure 6.
Spectral density function of mean weekly DO, Vermillion River, LA (Tabios 1984).

fitting a harmonic function (trigonometric series) to the means $\bar{X}_\tau$ and standard deviations S_τ by using harmonic or Fourier series analysis, and equation 8 is applied using such fitted functions. Both parametric and nonparametric methods have been widely used for seasonal standardization (see for instance Yevjevich 1972b, Salas et al. 1980, Hipel 1981).

The seasonal standardization derived by using equation 8 removes only the seasonality in the mean and variance as above stated. If the original series $X_{\nu,\tau}$ had other periodic components such as periodicity in the skewness, kurtosis and dependence structure, then the series $Z_{\nu,\tau}$ will remain with such periodicities. Removing periodic skewness or kurtosis may require nonlinear transformations based on the assumed marginal distribution corresponding to each season. Most of the available methods applied to analyzing and modeling water quality variables consider removing the skewness of the data so as to render a symmetrically (hopefully normally) distributed series. This can be done by using Box-Cox transformations or in many cases simply using log-transformations applied to the series $X_{\nu,\tau}$ as a whole (McLeod et al. 1983) or season by season (Salas et al. 1980). The log transformation can always be applied to variables that are known or assumed to be lognormally distributed and sometimes works well for other (moderately skewed) distributions. No similar transformations have been suggested to deal with kurtosis, although one can verify how the transformations used for skewness affect the kurtosis. It is expected that once the original data have been transformed and the skewness is within an acceptable range of zero, then the kurtosis will also be within an acceptable range of 3 (the theoretical value for a normal variable). This of course may not always be the case in actual applications.

Dependence Structure

Time dependence is one of the inherent statistical characteristics of most geophysical time series including water quality. In the case of nonpoint-source water quality series arising from agricultural activities, one may postulate that time dependence results from the natural travel of the water quality constituents through a series of storage reservoirs, particularly the unsaturated and the saturated zones. Thus, as with streamflow processes, water containing a certain constituent will be gradually released to a stream from a series of storage reservoirs (surface and subsurface) over a series of successive intervals, thus producing dependence in time. Quite likely, if most of the water quality constituents under study would reach the stream through surface drainage systems, then time dependence of the corresponding variables would not be significant.

A simple procedure for investigating the time dependence structure of water quality time series is correlogram analysis. Equation 6 gives the sample correlation of a given time series. It was noted previously that the correlogram of the seasonal variable $X_{\nu,\tau}$ (see fig. 5) yields a correlogram that is periodic because of the seasonality that may be present in the original series. Thus, one would tend to think that if $X_{\nu,\tau}$ is seasonally standardized by equation 8 (assuming also that the series was previously transformed into normal), then the remaining series $Z_{\nu,\tau}$ is free of cycles and is stationary for all purposes. In fact, this has been the typical assumption in most water quality analysis, modeling and applications thereof. For instance, the correlogram of the seasonally standardized series of monthly conductivity of the Yampa River is shown in figure 7. There is nothing unusual about this correlogram. In fact, by examining this correlogram one might assume that a simple lag-1 autoregressive model (eq. 10 defined later in this paper) would be a quite reasonable model for describing the correlation structure. If one analyzes the correlation structure more thoroughly, however, a different conclusion is reached.

Specifically, one may define the correlation structure of a given series season by season as

$$r_{k,\tau} = \frac{(1/N) \sum_{\nu=1}^{N} (X_{\nu,\tau} - \bar{X}_\tau)(X_{\nu,\tau-k} - \bar{X}_{\tau-k})}{S_\tau S_{\tau-k}} \qquad [9]$$

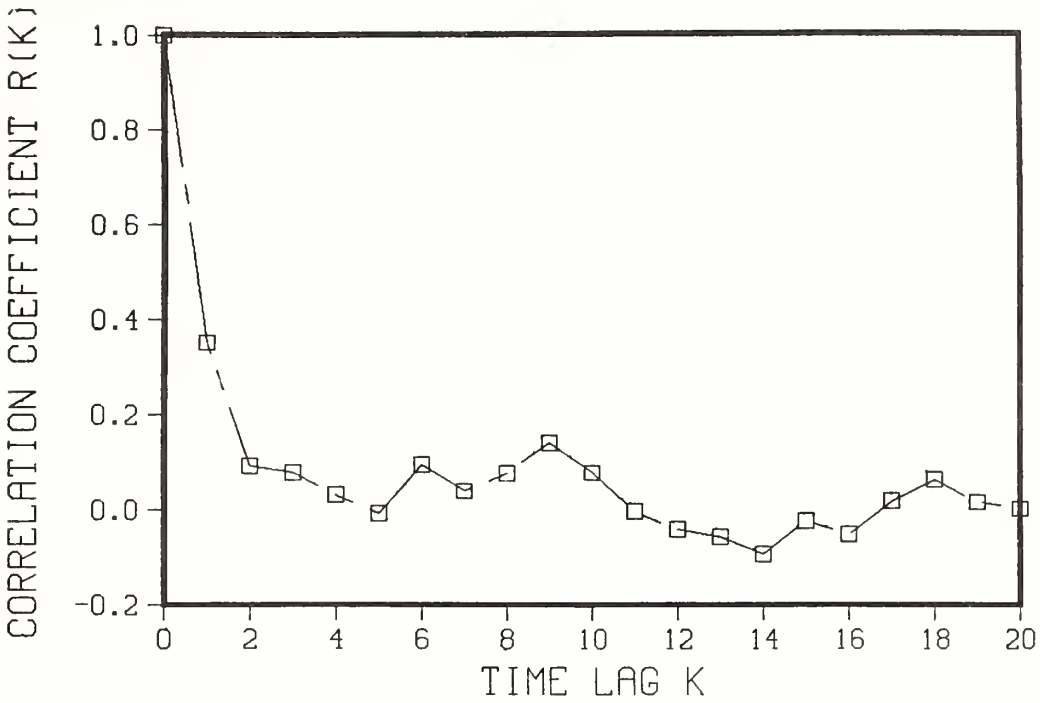


Figure 7.
Correlogram of equation 6 for the seasonally standardized series $Z_{\nu,\tau}$ of equation 8 based on monthly water conductivity, Yampa River, CO.

in which the mean $\bar{X}_\tau$ and standard deviation S_τ are as given by equations 2 and 3, respectively. In addition, if τ-k is less than or equal to zero it is replaced by τ-k+ω, N is replaced by N-1, and the index ν is replaced by ν-1. In equation 6, each value of r_k is estimated using all pairs of observations k time intervals apart. In equation 9 however, a separate value of r_k is computed for each season τ considering only observations for season τ and season τ-k. This way of defining the correlation structure has been widely used in statistical hydrology in general (Salas et al. 1980), but apparently it has not been used in most analyses of water quality time series.

As an example, figure 8 gives the lag-1 ($r_{1,\tau}$) and lag-2 ($r_{2,\tau}$) month-to-month correlations of monthly conductivity of Yampa River. The correlation structure is seasonal (cyclic or periodic) in the same sense as the seasonality of the other statistical characteristics such as the mean and variance. Likewise, figures 9 and 10 give another example of the overall correlogram r_k of the seasonally standardized series $Z_{\nu,\tau}$ and the seasonal correlogram $r_{k,\tau}$ of the original series for the monthly conductivity of Twin Lakes, Colorado (Twin 4). Theoretically, it may be shown (Salas and Smith 1980) that a process with a seasonal correlation structure, such as that shown in figure 8, will have a corresponding overall correlogram r_k (estimated by eq. 6) with a decreasing wavelike shape. Due to sampling variability, however, identifying such wave like correlograms from a plot such as figure 7 would be practically impossible.

Although emphasis thus far has been on dependence in time of a single series, equally important may be analyzing the dependence in time and space of the same water quality constituent or of two or more constituents. For instance, one may be interested in studying the space-time dependence of nitrate found in an aquifer sampled at several points or of nitrate and phosphorus found at a given point. Often water quality constituents are correlated to water flow. For instance, Lane (1975) found that daily specific conductivity and water discharge are seasonally correlated. Likewise, he found that ion or constituent proportions (ratio of its concentration to the total salt concentration) are related to discharge for a number of water quality variables such

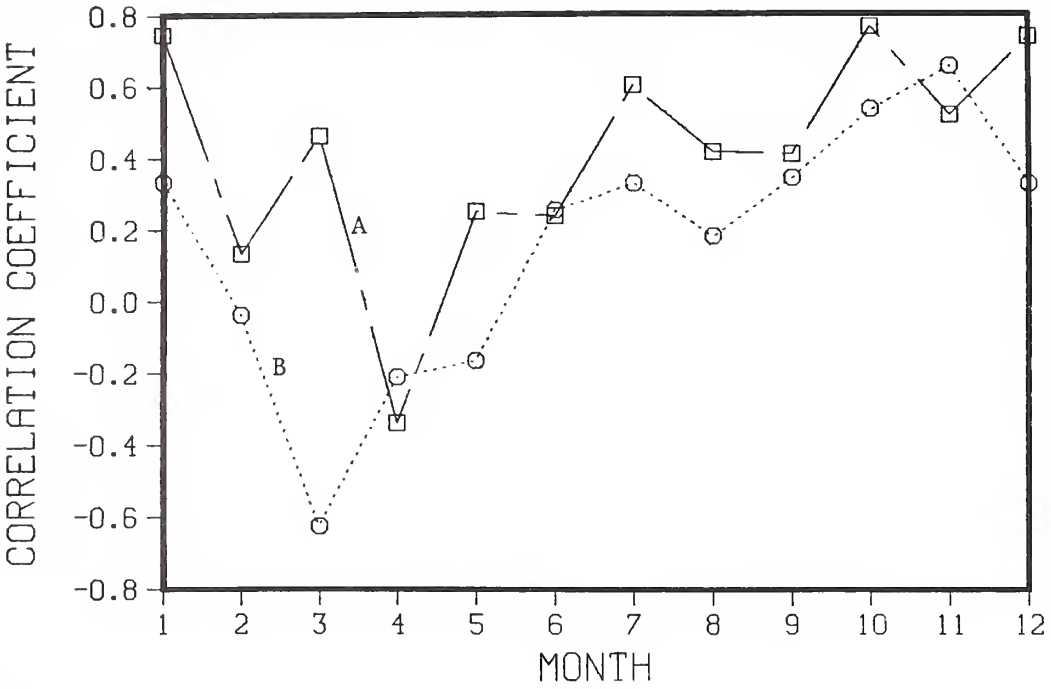


Figure 8.
(A) Lag-1 and (B) Lag-2 month-to-month correlations of monthly water conductivity, Yampa River, CO.

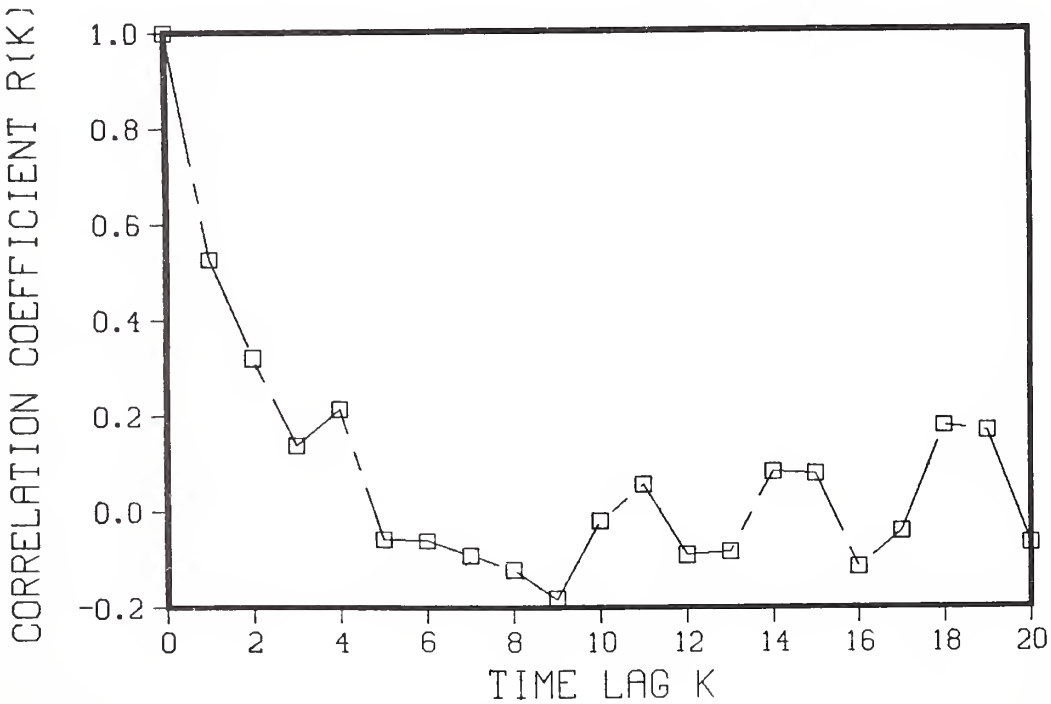


Figure 9.
Correlogram of equation 6 for the seasonally standardized series $Z_{\nu,\tau}$ of equation 7 based on monthly water conductivity, Twin Lakes, CO (Twin 4).

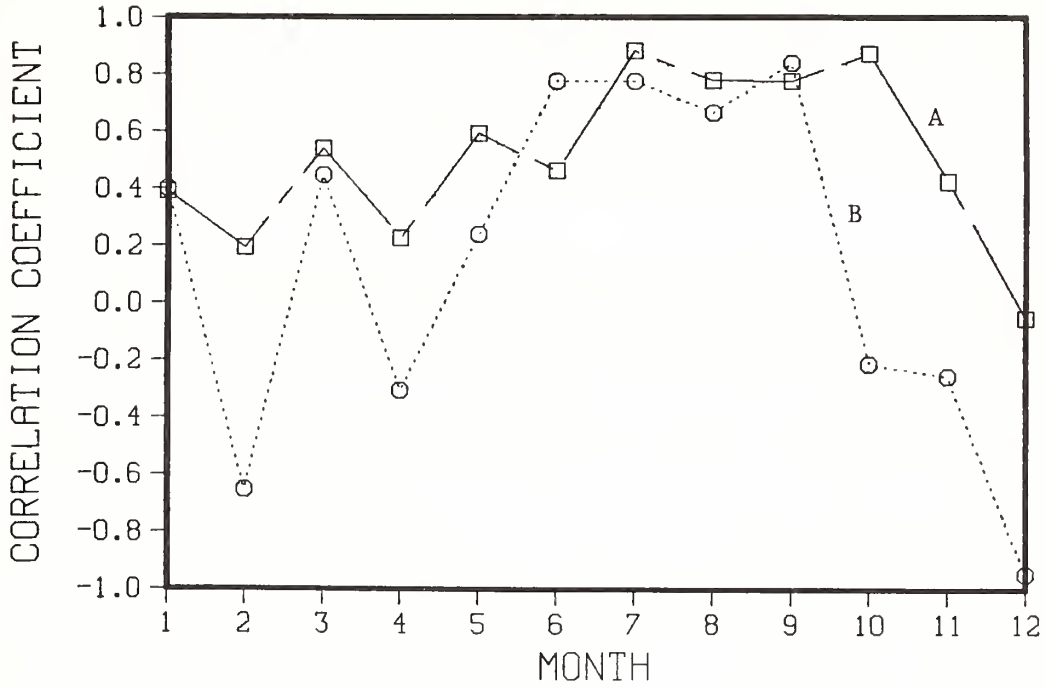


Figure 10.
(A) Lag-1 and (B) Lag-2 month-to-month correlations of monthly water conductivity, Twin Lakes, CO (Twin 4).

as calcium, sodium, potassium and sulfate. Such dependences may be estimated from discrete sample data by the usual cross-correlation analysis (Salas et al. 1980) or by cross-spectral analysis (Jenkins and Watts 1969).

MODELING OF WATER QUALITY SERIES

A number of approaches have been suggested for modeling water quality time series (Lee 1972, Litwin and Joeres 1976, D'Astous and Hipel 1979). Most of them follow the well-known Box-Jenkins approach (Box and Jenkins 1970) which essentially involves the following steps:

(1) Detrending the original series
(2) Transforming the remaining series into normal by log or other (Box-Cox) transformation
(3) Deseasonalizing the residual series
(4) Fitting the so-called autoregressive moving average models to the series resulting from step (3).

Step 1 will be specifically discussed in the following section in a broader context of detecting and modeling changes in water quality data. Steps 2 and 3 were discussed previously in section 2. In this section we elaborate further on step 4. In reality, we will simply refer to the models without going into much detail about properties, estimation, and overall fitting techniques.

Assuming that series $Z_{\nu,\tau}$ of equation 8 is second-order stationary (stationary in the mean and covariance), the simplest model is the first-order autoregressive defined as

$$Z_t = \phi_1 Z_{t-1} + \epsilon_t \qquad [10]$$

in which $Z_{\nu,\tau}$ has been redefined as simply Z_t with t in months, weeks or days or, in general, seasons, as the case may be; ϕ_1 is the autoregressive parameter; and ϵ_t is normal with mean zero and variance σ_ϵ^2. The autocorrelation function of this model decays exponentially. Including a second autoregressive term in equation 10 will lead to the second-order model. In general, a p-order model will include p autoregressive coefficients.

A further extension of equation 10 is the first order autoregressive moving-average model commonly known as an ARMA(1,1), which is given by

$$Z_t = \phi_1 Z_{t-1} + \epsilon_t - \theta_1 \epsilon_{t-1} \quad [11]$$

in which ϕ_1 is the autoregressive parameter, θ_1 is the moving average parameter and ϵ_t is normal with mean zero and variance σ_ϵ^2. Model 11 has a more "flexible" correlation structure than model 10, and as a result it is better suited to fit a wide variety of series. Including p autoregressive terms and q moving average terms will lead to the ARMA(p,q) model.

These models can be further extended to include differencing and "seasonality" (in the Box-Jenkins sense). For instance, the ARIMA (p,d,q) model is given by

$$\phi(B)(1-B)^d Z_t = \theta(B) \epsilon_t \quad [12]$$

where

$$\phi(B) = 1-\phi_1 B - \ldots - \phi_p B^p,$$

$$\theta(B) = 1-\theta_1 B - \ldots - \theta_q B^q,$$

and ϵ_t is normal with mean zero and variance σ_ϵ^2. In addition, B is the backward shift operator, i.e., $B^j Z_t = Z_{t-j}$. The above ARIMA does not have a "seasonal" component. Thus, the general ARIMA model is written as

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D Z_t = \theta(B)\,\Theta(B^s)\,\epsilon_t \quad [13]$$

where s is the seasonal period (i.e., s = 12 for monthly data), $\Phi(\cdot)$ is a polynomial of degree P and $\Phi(\cdot)$ is a polynomial of degree Q. In both models the nonstationary part is in the $(1-B)^d$ and/or $(1-B^s)^D$ term. For example, a seasonal stationary model arises if d = D = 0. The usual assumptions regarding the roots of the polynomials are understood (Box and Jenkins 1970).

A number of books and papers have been written applying the foregoing models to geophysical time series in general (Salas et al. 1980, Loucks et al. 1981, Bras and Rodriguez-Iturbe 1985) and to water quality variables in particular (Lettenmaier 1976, Sanders and Adrian 1978, D'Astous and Hipel 1979, Loftis and Ward 1980a, McLeod et al. 1983).

As illustrated in the previous section some water quality time series may show periodic dependence structure. If so, periodic autoregressive (PAR) and periodic autoregressive moving average (PARMA) models may be useful. The PAR(1) model has been widely used in stochastic hydrology (Hannan 1955, Thomas and Fiering 1962, Yevjevich 1972a, Salas et al. 1980, Loucks et al. 1981, Bras and Rodriguez-Iturbe 1985). The PARMA(1,1) model has been suggested for modeling seasonal hydrologic time with complex dependence structure (see for instance Tao and Delleur 1976, Hirsch 1979, Salas et al 1980, Salas et al. 1982). Both PAR(1) and PARMA(1,1) models may be useful for modeling water quality time series exhibiting complex periodic dependence structure.

The PAR(1) model may be written as

$$Z_{\nu,\tau} = \phi_\tau Z_{\nu,\tau-1} + \epsilon_{\nu,\tau} \quad [14]$$

in which ϕ_τ is the periodic autoregressive parameter and $\epsilon_{\nu,\tau}$ is normal with mean zero and periodic variance $\sigma_\tau^2(\epsilon)$. Based on the method of moments it may be shown that ϕ_τ may be estimated by $r_{1,\tau}$ of equation 9 and the noise variance is $1-r_{1,\tau}^2$. Thus, the model is very simple to apply. Of course, as in the case of model 10, extensions of model 14 to higher orders is possible (Salas et al. 1980).

Likewise, the PARMA(1,1) model may be written as

$$Z_{\nu,\tau} = \phi_\tau Z_{\nu,\tau-1} + \epsilon_{\nu,\tau} - \theta_\tau \epsilon_{\nu,\tau-1} \quad [15]$$

in which ϕ_τ and θ_τ are the periodic autoregressive and moving average parameters, respectively, and $\epsilon_{\nu,\tau}$ is the normal noise with mean zero and periodic variance $\sigma_\tau^2(\epsilon)$. Estimation and some extensions of this model may be found in Salas et al. (1982).

Parameter estimation of the foregoing stationary and periodic models is well known (Fiering and Jackson 1971, Salas et al. 1980). Moment and maximum likelihood approaches are generally utilized. Moment estimators of low-order models are simple to obtain. Finding maximum likelihood estimators in general is somewhat more complex; however, algorithms and software for doing so are available (Box and Jenkins 1970, Brockwell and Davis 1986). Once model parameters are estimated, it is necessary to test whether the assumptions of the given model are met, i.e., normality, independence and homoscedasticity of the residuals. If the assumptions are not met, then a different model order is selected and tested until an appropriate model is found. In this process, further criteria may be useful for model selection. For instance, the Akaike information criteria (AIC) have been used for selecting the appropriate model when modeling water resources and other geophysical time series (Hipel 1981). Furthermore, additional model testing may depend on the purpose of the model. For instance, if the model is to be used for forecasting, then forecast comparisons are made; i.e., for a fraction of the data, a forecast is made and the results are compared with the actual historic observations (Litwin and Joeres 1976).

DETECTION AND MODELING OF CHANGES IN WATER QUALITY DATA

We said earlier that changes in nonpoint-source water-quality series may occur in several forms, particularly as trends and jumps or shifts. First, the phenomenon under study might have a cyclic (usually seasonal) component or not (such as the usual annual data). Within each cyclic or noncyclic component, one can consider single series or multiple series (for instance when sampling along a stream or at several points in an aquifer). Furthermore, the modeling of the phenomenon, which is analyzed for detecting changes, may include dependence structure or not. For the most part, realism suggests that some dependence structure is needed in water quality modeling; however, most reliable statistical results are for the independent case. There are still other ways of subcategorizing the problem. For instance, the time of change may be known or not. The type of change may be a change in the level (such as the mean), a change in variability (such as variance), or a change in dependence structure (such as serial correlation), or any combination of these and more. Theoretical results, for the most part, concentrate on change in level or mean. When considering two or multiple series, variations of the problem of detecting changes arise. For instance, one might test the equivalence of two or multiple series, or check for a change in one series given no change in the other, etc.

Although the topic suggests that interest is centered on the testing of hypotheses mode of statistical inference, the other two forms, point estimation and interval estimation, of classical

statistical inference are also pertinent. For instance, a point estimate or interval estimate of the amount of change in the mean level might be of interest. Also, the type of statistical analysis to be performed provides still another breakdown of the problem. One might use a statistical analysis based on the likelihood principle, or use Bayesian analysis, or some nonparametric approach. The entire problem cannot be covered in a single paper. Instead one or two topics will be selected and emphasized.

For those problems in which the time of change is known, the problem is often labeled intervention analysis. See, for instance, Box and Tiao (1975), Hipel et al. (1975) or Lettenmaier (1977). Intervention analysis is one topic that will be emphasized. And, in general, single series will be emphasized over multiple series. The statistical literature on the subject is vast, with several very recent publications. The book by Sanders et al. (1983) and that by Gilbert (1987) are excellent sources on some of these topics.

Detection and Modeling of Trends

There are two important aspects in analyzing long-term tendency: detection and characterization. There are a number of statistical methods for detecting time series tendencies. The biggest impediment to this task is the shortness of data records. It is still of current interest how the detected tendencies compare in significance with the relatively high standard errors of small samples as well as with typical measurement errors in monitoring. The problem of characterization of a detected tendency (linear or nonlinear, transient or nontransient, etc.) may be also a complex task, although simple techniques, such as inspection of time series plots of raw data, spectral analysis and linear regression analysis are widely used.

The observed raw data must be plotted against time before any time series modeling is started. Time series plots usually provide valuable information on the behavior and structure of the data. Among the time series components, long-term tendencies and periodicities are commonly or distinctly exhibited on these plots. Although time series plots provide essentially qualitative information, they are often helpful in identifying the mathematical structure of these components. A commonly applied test for detecting long-term trends is correlation analysis. For instance, the Spearman's rank correlation coefficient can be used as test statistic to test the hypothesis that there is no association between two populations. If time is one population and a given water quality variable arranged in time is another population, one can assess for an apparent trend over time. Many water quality variables have been tested for trends by using this technique (Jones et al. 1984). Likewise, spectral analysis may be useful for detecting long-term trends. An analysis of the variance spectrum may reveal trends, periodicities, and the degree of dependence that are not easily discernible in a time series. Specifically, the plot of the spectral density function against frequency may indicate a long-term tendency if the highest peak is at zero frequency. For instance, Tabios (1984) used this technique for identifying trends on a number of dissolved oxygen time series.

The use of simple linear regression between a given water quality variable and time has been widely used for detecting and modeling trends. In equation form, the linear trend is written as:

$$T_t = a + bt \qquad [16]$$

where a and b are coefficients obtained from regressing the raw observed data against time. If the slope b is not significantly different from zero, then there is no significant evidence of a linear trend. A statistical test for the significance of b may be found in any standard statistics book (see for instance Benjamin and Cornell, 1970). The t-test is commonly used for this purpose. However, it can give misleading results if periodicities and time dependence are present. Actually the same is true in applying Spearman's correlation test. Yevjevich (1972c) suggested that statistical tests

developed for uncorrelated data may be applied to correlated data if the equivalent independent sample size, as developed by Bayley and Hammersley (1946), is used. Lettenmaier (1976) advocates the applicability of equivalent independent sample size for testing changes in the mean and linear trends of correlated water quality samples. In addition, results reported by Hirsch et al. (1982) suggest that in cases of seasonality and dependence the seasonal Kendall test is preferred.

Detecting and Modeling Other Changes

In this section the emphasis is on detecting and modeling sudden or gradual changes in the mean. We closely follow an article written by Salas and Boes (1980). Assume that data are recorded in discrete time and the observations are available for N successive time epochs. Let $X_1,...,X_N$ denote the observed water quality data. Write $X_1,...,X_N$ as $X_1,...,X_{\tau-1},X_\tau,X_{\tau+1},...,X_N$, where it is assumed that a "change," which may or may not affect the distribution of the X_i's, occurs between time epochs τ and $\tau+1$, where $\tau=1,...,N-1$. One problem of detection-of-change is the following: assume $X_1,...,X_\tau$ have marginal cumulative distribution function (cdf) $F(\cdot)$ and $X_{\tau+1},...,X_N$ have marginal cdf $G(\cdot)$; test the null hypothesis that F is identical to G, that is, there is no change in marginal distribution before and after time epoch τ. If τ is assumed known and the X_i's are assumed independent, this reduces to the classical two-sample problem for which there are many standard statistical analyses. In fact, there are several nonparametric techniques that work quite well, such as the Kolmogorov-Smirnov two-sample test. For reference see any text on nonparametric statistics.

Case of Independence--Time of Change Known.
Suppose $X_1,...,X_\tau$ are independent and identically distributed (iid) normal random variables with mean μ_1 and variance σ^2 and $X_{\tau+1},...,X_N$ are iid normal with mean μ_2 and variance σ^2. Further, assume $X_1,...,X_\tau$ independent of $X_{\tau+1},...,X_N$. Then the standard two-sample t-test can be used to test the equality of μ_1 and μ_2. Let

$$t = \left[\frac{(\bar{X}_2-\bar{X}_1) - (\mu_1-\mu_2)}{\sqrt{(1-\tau)+[1/(N-\tau)]}}\right]\left[\frac{\sum_{1}^{\tau}(X_j-\bar{X}_1)^2+\sum_{\tau+1}^{N}(X_j-\bar{X}_2)^2}{N-2}\right]^{-1/2} \qquad [17]$$

where $\bar{X}_1 = (1/\tau)\sum_{1}^{\tau} X_j$ and $\bar{X}_2 = [1/(N-\tau)]\sum_{\tau+1}^{N} X_j$, then t has a t-distribution with N-2 degrees of freedom and with $\mu_1 = \mu_2$ serves as a test statistic for testing H_o: $\mu_1 = \mu_2$ versus H_1: $\mu_1 \neq \mu_2$ or H_1: $\mu_1 < \mu_2$. The usual results on the two sample t-tests apply. For example, the power of the test can be obtained and the magnitude of τ and N necessary to achieve a prescribed power for some fixed $\mu_2-\mu_1$ can be found. Also, t of equation 17 serves as a pivotal quantity (Mood et al. 1974) and can be used to give a confidence interval estimator of $\mu_2 - \mu_1$, which is the change in the mean level. This is all standard statistical theory and will not be discussed further.

A similar setup leads to the F-test as a vehicle for detecting change in the variance. Suppose now that $X_1,...,X_\tau$ are iid $N(\mu_1,\sigma_1^2)$ and $X_{\tau+1},...,X_N$ are iid $N(\mu_2,\sigma_2^2)$ and $X_1,...,X_\tau$ are independent of $X_{\tau+1},...,X_N$. To test H_o: $\sigma_1^2=\sigma_2^2$, the test statistic

$$f = \left[\sum_{1}^{\tau}(X_j-\bar{X}_1)^2/(\tau-1)\right]\left[\sum_{\tau+1}^{N}(X_j-\bar{X}_2)^2/(N-\tau-1)\right]^{-1} \qquad [18]$$

can be used; under $\sigma_1^2=\sigma_2^2$ it has an F-distribution with τ-1 and N-τ-1 degrees of freedom, and gives the standard F-test for testing equality of variances. One could also obtain a confidence interval estimate for the ratio of variances which would indicate the magnitude of change in variance.

The two sample t-test for testing equality of means and the F-test for testing equality of variances require an assumption of normality. There are nonparametric techniques for testing the equality of location parameters (such as means) and the equality of scale parameters (such as standard deviations) that do not require the normality assumption, but these will not be discussed. If the normality assumption is retained and one assumes several known change epochs (rather than just one), one can readily generalize from testing the equality of two means to testing the equality of several means. The resulting test is that of a one-way analysis variance (see for instance Matalas and Dawdy 1964). Likewise, one can generalize from testing the equality of two variances to testing the equality of several variances. For modifications of the t- and F-tests for the case of correlated water quality data see Lettenmaier (1976).

Case of Independence--Time of Change Unknown

In the previous subsection we assumed that a change occurred between times τ and $\tau+1$. Now we address the case where the time of change is not known. This case has attracted considerable attention in the statistical literature (see for instance Hinkley 1970, Sen and Srivastava 1975a and 1975b, Lee and Heghinian 1977). Consider the following setup (Lee and Heghinian 1977):

$$X_j = \begin{cases} \mu + \epsilon_j & , \quad j = 1,2,...,\tau \\ \mu + \delta + \epsilon_j & , \quad j = \tau+1,...,N \end{cases} \qquad [19]$$

where the ϵ_j are iid normally distributed with mean zero and variance σ^2, and τ, μ, δ, and σ are all unknown parameters. They derived the joint and marginal posterior distributions of τ and δ, the point of change and amount of change, respectively. The following priors are used: τ is assumed uniform over 1,2,...,N-1, δ and μ are each assumed normal with zero means but different variances and σ is assumed distributed proportional to the inverse of σ; and all are assumed independent. Also $\sigma^2/N << \sigma_\mu^2$ and $\sigma^2 << \sigma_\delta^2$. The joint posterior distribution of τ and δ is given by

$$f(\tau,\delta \mid X_1 = x_1,...,X_N = x_N) \alpha$$
$$\{H(\tau) + [\tau(N-\tau)(\delta - \hat{\delta}_\tau)^2]/N\}^{-(N-1)/2} \qquad [20]$$

where α denotes "is proportional to", $-\infty < \delta < \infty$ and $\tau = 1,...,N-1$, where

$$H(\tau) = \sum_1^\tau (x_i-\bar{x}_\tau)^2 + \sum_{\tau+1}^N (x_i-\bar{x}_{N-\tau})^2$$

and

$$\hat{\delta}_\tau = \bar{x}_{N-\tau} - \bar{x}_\tau = [1/(N-\tau)] \sum_{\tau+1}^N x_i + (1/\tau) \sum_1^\tau x_i$$

From the joint posterior, the marginal posteriors are readily obtained. On the basis of these posterior distributions one can readily make inferences regarding τ and δ. The posterior distribution of τ is given by

$$f(\tau \mid X_1 = x_1,\ldots,X_N = x_N) \propto$$

$$\{N/(\tau(N-\tau))\}^{1/2} \left\{ \frac{\sum_{i=1}^{\tau} (x_i-\bar{x}_\tau)^2 + \sum_{i=\tau+1}^{N} (x_i-\bar{x}_{N-\tau})^2}{\sum_{i=1}^{N} (x_i-\bar{x})^2} \right\}^{-(N-2)/2} \qquad [21]$$

for $\tau = 1,2,\ldots,N-1$

Similarly, the posterior distribution of the amount of change is given by

$$f(\delta \mid X_1 = x_1,\ldots,X_N = x_N) \propto \sum_{\tau=1}^{N-1} f(\tau,\delta \mid X_1 = x_1,\ldots,X_N = x_N) \qquad [22]$$

Salas and Boes (1980) applied this approach for testing changes in precipitation. Although this approach is somewhat more complex than the t- and F-tests, it gives more statistical information for detecting and quantifying data changes. The foregoing approach may be useful for testing changes of water quality time series when there is no certainty as to when the change may have occurred. This may be specially attractive in case of nonpoint-source water-quality variables in which time of changes may be difficult to know because of the response delay and the unknown source(s) of the constituents.

Case of Dependence--Intervention Analysis

In all of the preceding, we assumed independence of the X_j's. In practice not only are successive observations usually dependent, but at least occasionally the time series would be nonstationary. A type of stochastic model that allows for dependence as well as a mild form of nonstationarity is the so-called ARIMA model (Box and Jenkins 1970).

Box and Tiao (1965) were the first to attempt to detect a change in level of a nonstationary dependent time series. They considered an integrated moving average process. After this first attempt the term "intervention analysis" caught on and was considered by many authors. See Box and Tiao (1975), Hipel et al. (1975), Hipel et al. (1977), Lettenmaier (1976) and D'Astous and Hipel (1979). Intervention analysis is a stochastic modeling technique designed to determine whether or not a natural or artificially-induced intervention causes a significant change in the level of a time series. Its applications in water resources are extensive (Hipel et al. 1975).

The intervention analysis model as described here can handle a seasonal component. As the term suggests, the times of "intervention" are usually known, so in the language of the above τ is known, although our description is different. Intervention analysis involves building a model in the usual Box-Jenkins terminology. That is, one follows the iterative scheme of model identification, model fitting and diagnostic checking. We follow Box and Tiao (1975) and Hipel et al. (1975).

Let $x_1,\ldots,x_N$ represent observations of $X_1,\ldots,X_N$. We model $X_1,\ldots,X_N$ by modeling $\ldots,X_{j-1},X_j,X_{j+1},\ldots$. Let Y_j be some appropriate transformation (for example, a log or power transformation) of X_j and work with the Y_j's. Assume

$$Y_t = d(\kappa,\xi,t) + Z_t,\ t=\ldots-1,0,1,\ldots \qquad [23]$$

where $d(\kappa,\xi,t)$ is the nonstochastic dynamic part of the model including the intervention component, and Z_t is the stochastic part of the model or "noise" term. The term $d(\kappa,\xi,t)$ can include deterministic effects, κ is a set of parameters indexing $d(\cdot,\cdot,\cdot)$, ξ is a set of exogenous

variables, and t is the time index. In addition, the term Z_t is modeled by an ARIMA process as in equation 13.

The dynamic nonstochastic part of the model with general notation $d(\cdot,\cdot,\cdot)$ may be clarified with some examples. Suppose, for instance, it is desired to represent only a step change of magnitude ω between times τ and $\tau+1$, i.e.,

$$Y_t = \begin{cases} Z_t & \text{for } t = \ldots,\tau-1,\tau \\ \omega + Z_t & \text{for } t = \tau+1,\ldots \end{cases} \qquad [24]$$

then $d(\kappa,\xi,t) = \omega\, I_{\{\tau+1,\tau+2,\text{--}\}}(t)$, where

$$I_A(t) = \begin{cases} 1 \text{ if } t \text{ belongs to } A \\ 0 \text{ otherwise} \end{cases}$$

is the usual indicator function.

More generally, suppose there is a single exogenous variable, say ξ_t. Suppose that the transfer, Y_t, to the output from the input ξ_t is generated by the linear difference equation

$$\delta(B)\, Y'_t = \omega(B)\, \xi_t \qquad [25]$$

where $\delta(\cdot)$ and $\omega(\cdot)$ are polynomials similar to $\phi(\cdot)$ or $\theta(\cdot)$ of equation 12, i.e.,

$$\delta(B) = 1-\delta_1 B - \ldots -\delta_r B^r$$

$$\omega(B) = \omega_0-\omega_1 B - \ldots -\omega_s B^s$$

Then

$$Y_t = Y'_t + Z_t = (\omega(B)/\delta(B))\, \xi_t + Z_t \cdot \qquad [26]$$

If $\xi_t = I_{\{\tau,\tau+1,\ldots\}}(t)$, $\delta(B) = 1$, and $\omega(B) = \omega B$ then equation 26 reduces to equation 24. Thus equation 25 is a convenient method of representing the effect of a single exogenous variable ξ_t. The most general representation of the dynamic nonstochastic function $d(\kappa,\xi,t)$ that we will consider is

$$d(\kappa,\xi,t) = \sum_{j=1}^{k} Y_{tj} = \sum_{j=1}^{k} \{\omega_j(B)/\delta_j(B)\}\, \xi_{tj} \qquad [27]$$

where $\{\xi_{t1}\},\ldots,(\xi_{tk})$ represent k different exogenous variables (with a potential of k different changes). Here the parameters κ are the coefficients in the polynomials $\omega_j(B)$'s and $\delta_j(B)$'s. This is actually quite powerful representation for the dynamic component of $\{Y_t\}$ and is patterned after the modeling of the stochastic component. Suppose, for instance, there is a single intervention that takes effect gradually. The following Y'_t would do:

$$\frac{\omega B}{1-\delta B} I_{\{\tau,\tau+1,\ldots\}}(t)$$

with parameters ω and δ. Note that Y'_t is given by

$$\omega B(1+\delta B + \delta^2 B^2 + \ldots)\, I_{\{\tau,\tau+1,\ldots\}}(t)$$

$$= \omega(B + \delta B^2 + \delta^2 B^3 + \ldots)\, I_{\{\tau,\tau+1,\ldots\}}(t)$$

$$= \begin{cases} 0 & \text{for } t = \tau, \tau-1, \tau-2, \ldots \\ \omega & \text{for } t = \tau+1 \\ \omega(1+\delta) & \text{for } t = \tau+2 \\ \omega(1+\delta+\delta^2) & \text{for } t = \tau+3 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{cases} \qquad [28]$$

which gives an immediate change of ω and an ultimate change of $\omega/(1-\delta)$.

A procedure to use in applying the above is as follows: For a transform of the data from $x_1,\ldots,x_N$ to $y_1,\ldots,y_N$, restrict to models of the form

$$Y_t = \sum_{j=1}^{k} Y'_{tj} + Z_t. \qquad [29]$$

Decide from physical considerations how many exogenous variables and how many interventions there are and the effect of the interventions. That is, identify the dynamic nonstochastic part of the model. Suppose, for illustrative purposes, that there is a single exogenous variable represented by an intervention between τ and $\tau+1$ that has a gradual effect that we assume is of the form of equation 28. We next need to identify the seasonal ARIMA model for the noise term Z_t. Proceed as in Box and Jenkins (1970). That is, identify a particular ARIMA model by specifying p, q, P, Q, d, and D. Again, for illustrative purposes, suppose we try a nonseasonal stationary ARMA(1,1) model, i.e., P = Q = D = d = 0 and p = q = 1. We have parameters ϕ, θ, and σ_ϵ^2 in our stochastic model. Now estimate the parameters of both the dynamic part and the stochastic part by the usual nonlinear least squares routine. For our illustrative example we get estimates of ω, δ, ϕ, θ and σ_ϵ^2. This estimating of the parameters is fitting the model. Next do a diagnostic check by analyzing the residuals. If the check fails, modify the model accordingly and repeat.

Box and Tiao (1975) gave two examples, one dealing with photochemical smog in Los Angeles and the other with changes in the consumer price index surrounding the Nixon Administration's institution of Phase I and II of wage-price controls. Both sets of data were monthly and included a seasonal component. Hipel et al. (1975) gave a nonseasonal example using annual flows of the Nile River surrounding the building of the Aswan Dam and D'Astous and Hipel (1979) gave another example using phosphorus data downstream from the Guelph sewage treatment plant, Speed River, Ontario. The technique shows up well in all examples. In summary, although the technique is not entirely objective and is computationally demanding, the so-called interventional analysis model appears to be quite useful in detecting changes in hydrologic data. Any improvements in the so-called Box-Jenkins modeling will result in companion improvements in the intervention analysis. One must acknowledge that the significant data requirements of the model fitting process will limit the applicability of intervention analysis in studying nonpoint-source pollution. There is a potential, though, for application in intensively studied regions such as the Chesapeake Bay or the San Joaquin Valley.

DESIGN OF SAMPLING PROGRAMS

A water quality sampling program dealing with nonpoint-source pollution can logically be designed based on stochastic models, such as those we have discussed, of the water quality

processes of interest. Usually we have one or two major goals in mind, the most common being the estimation of annual loads of certain constituents or detection of trends in loads. The first step in sampling program design should generally be the formulation of monitoring objectives in statistical terms. Examples of statistical objectives follow:

(1) Estimate the mean annual phosphate loading to Cherry Creek Reservoir with a 90% confidence interval width no greater than 20% of the mean.

(2) Detect any changes in annual mean phosphate loading to Cherry Creek Reservoir greater than 10% of the current loading with a probability of detection or power of 80% over a 5-year sampling horizon.

The required sampling locations and frequencies can be found once such concrete objective statements are formulated, providing that one has an adequate model of the process being monitored. But of course, one cannot have a good stochastic model unless a long data record (greater than 5 years of monthly data) is available from a sampling program already in place. Thus detailed stochastic process models are not of much use in designing new sampling programs, although they may sometimes be used to improve existing monitoring activities.

To start a new monitoring effort one can often use rules of thumb, (Sanders et al. 1983, Loftis and Ward 1980a, and Loftis and Ward 1987) developed from the experience of practitioners, to design a limited preliminary or first-stage monitoring program. The results of this program would be used to refine the monitoring network, gradually adding stations over time. Alternatively, one can use a short-term intensive survey to formulate statistical models which can then be used as a basis for design of long-term monitoring. The latter approach, although meritorious, is rarely used due to lack of funds to perform the intensive survey or due to lack of patience and a desire to get going on the long-term monitoring.

Design Criteria for Estimating Loads and Average Concentrations

The number of samples required to estimate average conditions, such as mean annual loadings or mean concentration of a certain constituent, may be partially based on the desired width of the confidence interval about the mean, as suggested by the objective statement in the Cherry Creek Reservoir example. The confidence interval about a sample mean may be expressed as follows (Snedecor and Cochran 1980).

$$\bar{X} \pm t_{\alpha} \frac{S}{\sqrt{N}} \qquad [30]$$

where $\bar{X}$ = sample mean
S = sample standard deviation
N = number of observations used to compute $\bar{X}$ and S
t_{α} = Student's t statistic with two-tailed exceedence probability α and N-1 degrees of freedom.

Although the above expression relies on an assumption of normality of the sample mean, the Central Limit Theorem implies that the sample mean will tend to be normally distributed, regardless of the distribution of the individual observations. However, if the distribution is very skewed to the right, one may wish to work with the geometric mean or median instead of the arithmetic mean; expressions for the appropriate confidence interval may be found in Gilbert (1987).

In practice, one seldom specifies a desired confidence interval width right from the start. Rather one considers the range of feasible sample sizes or sampling frequencies and then computes

corresponding confidence interval widths. The final choice of sample size is a compromise between precision of estimates and cost of sampling.

Two major factors complicate sampling design using the confidence interval approach. The first and most obvious is that loads of nonpoint-source constituents are often largely contributed by storm events which occur stochastically. Thus a uniform sampling frequency will be of little use for estimating mean annual loads. One could, however, use the confidence interval approach to determine the number of storm events to monitor in order to estimate long-term average conditions, assuming that each event represents an independent observation.

Within each monitored event, continuous monitoring of flow is generally feasible. One may use grab samples for measuring water quality variables during the storm event at specified intervals of time or flow. In some cases, flow vs. quality regressions (from previous intensive surveys) may be used to make loading estimates in the absence of water quality measurements for a given event or to improve loading estimates based on sparse water quality measurements. Note, however, that bias concerns are important when using log-log regression equations (Koch and Smillie 1986).

The precision of the estimated load for a given storm event may be somewhat difficult to determine. However, if the constituent loading from an individual storm event is treated as a single observation and the observed storms are selected at random, the above expression for confidence interval width will be correct for long term average conditions. Estimates may be improved, however, by using a stratified random sampling approach rather than selecting storms purely at random (Gilbert 1987). In this case one would divide storms into classes, perhaps according to type or magnitude, and then sample a few storms randomly from each type. To effectively implement such a strategy would of course require long term monitoring of selected watersheds. To accurately reflect regional conditions, monitoring locations (watersheds) should be chosen by a stratified random sampling approach as well.

The second complicating factor associated with the confidence interval approach is the presence of serial correlation, which violates the independence assumption of equation 30 above. If long data records are available for the location of interest, one can employ ARIMA modeling or other approaches to characterize the temporal correlation structure and compute corresponding confidence interval widths (Loftis and Ward 1980a). A more widely applicable approach, however, is to average observations over time (or flow) intervals in such a way that the resulting averages may be regarded as independent. The above approach dealing with loading from individual storms is an example.

Design Criteria for Detecting Trends

The required sampling frequency for detecting changes in water quality of a given magnitude with a given level of certainty may be determined from the power of an appropriate statistical test for trend. The power of a hypothesis test for trend is the probability of concluding that a trend is present when one is really there. Power is directly related to the trend magnitude and inversely related to the confidence level of the test. The confidence level is defined as the probability of concluding that no trend is present when in fact there is none.

In general, the trend is best detected when the observations are equally spaced in time. However, as we mentioned above, each observation may be obtained by averaging or transforming in some way a large number of field measurements.

When the individual observations may be regarded as independent and the errors are normally distributed, the t-test for significance of regression slope will be an appropriate test for linear trend. The following approximate expressions from Lettenmaier (1976) may then be used to relate sampling frequency to the power of trend detection. Let us define a trend number N_t as follows:

$$N_t = \frac{\tau\sqrt{[N\ (N+1)(N-1)]}}{\sqrt{12}\ \sigma} \qquad [31]$$

where N = number of observations used in regression
τ = trend magnitude in units per sampling interval
σ = standard deviation of water quality variable in absence of trend and seasonality

Then power of the test is given by $1-\beta=F(N_t - t_{1-\alpha/2,\nu})$, where F is the cumulative distribution function of the Student's "t" distribution with $\nu=N-2$ degrees of freedom (here $\alpha/2$ is a one-tailed exceedence probability).

In most real cases, one should not be comfortable with assumptions of normality or independence, even at the preliminary design stage of a new sampling program. However, the above equation can still be used as a guide, provided that one can roughly estimate the number of independent samples per year which can be obtained from the system of interest. The linear regression model can also serve as the basis for design when one wishes to estimate the magnitude of a trend as well as to test for significance. In this case a confidence interval width around the estimated trend magnitude or slope may be used as a design criterion.

For estimating the obtainable effective independent sample size when no background data are available, one might start with the following extremely rough rule of thumb. Often about 4 independent samples per year can be obtained in groundwater or lake quality monitoring, and about 6 to 12 independent samples per year can be obtained in stream quality monitoring (Loftis and Ward 1980a, and Loftis et al 1986). In general, the more stable the flow pattern of interest, the greater the degree of serial correlation in both flow and quality, thus the smaller the number of independent samples that can be obtained.

In many cases it may be necessary or desirable to sample at a frequency greater than that which produces independent samples. For example, a monitoring objective of detecting extreme events would usually require frequent, perhaps continuous, sampling. Furthermore, estimation and trend detection ability are improved by more frequent sampling, even if some of the "added information" is redundant. For example, consider two series, one consisting of quarterly individual observations and the other of quarterly values each of which is the average of three monthly observations. Assume both series contain four independent samples. However, all else being equal, the averaged series will have a lower variance and thus contain more information.

A similar approach may be followed if one is interested in detecting step trends instead of gradual (linear) trends. In this case the power of the two-sample t- test may be used as a starting point (Snedecor and Cochran 1980).

Once the monitoring program is under way, initial assumptions regarding statistical characteristics of data can be refined, and appropriate tests for trend can be selected. For routine use, especially in relatively "young" sampling programs, we recommend the use of nonparametric methods which account for seasonal patterns in the mean. The seasonal Kendall test of Hirsch et al. (1982) has already been mentioned, and Gilbert (1987) provided a comprehensive discussion of this and other practical approaches.

As we mentioned for estimating average conditions, one rarely specifies the required statistical performance of a sampling program in the form of rigid design criteria. Rather one usually selects a sampling frequency by considering a range of feasible frequencies and evaluating the tradeoffs between power of trend detection and cost of sampling. When a sampling program is designed to both estimate average conditions and detect trends, one must also compromise sampling frequencies in order to accommodate both objectives within a reasonable budget.

ACKNOWLEDGMENTS

We would like to acknowledge the Colorado Agricultural Experiment Station projects "Characterizing the Properties of Precipitation for Agricultural Areas in Colorado" N. 0357 and "Maintenance and Protection of Water Quality for Agriculture" N. 0349, for the support in writing this paper.

LITERATURE CITED

Bayley, G.V., and J.M. Hammersley. 1946. The effective number of independent observations in an autocorrelated time series. J.R. Statist. Soc. 8:184-197.

Beck, M.B. 1987. Water quality modeling: A review of the analysis of uncertainty. Water Resour. Res. 23:1393-1442.

Benjamin, J.R., and C.A. Cornell. 1970. Probability statistics and decision for civil engineers. McGraw-Hill, New York.

Box, G.E.P., and G.C. Tiao. 1965. A change in level of a non-stationary time series. Biometrika 52:181-192.

Box, G.E.P., and G.C. Tiao. 1975. Intervention analysis with applications to economic and environmental problems. J. Am. Statist. Assoc. 70:70-79.

Box, G.E.P., and G.M. Jenkins. 1970. Time series analysis: forecasting and control. 1st ed. Holden-Day, San Francisco.

Bras, R., and I. Rodriguez-Iturbe. 1985. Random functions and hydrology. Addison-Wesley, Reading, MA.

Brockwell, P.J., and R.A. Davis. 1986. Time series theory and methods. Springer, Berlin.

D'Astous, F., and K.W. Hipel. 1979. Analyzing environmental time series. J. Environ. Eng. Div. ASCE 105:979-992.

Fiering, M.B., and B. Jackson. 1971. Synthetic hydrology. Monograph No. 1, American Geophysical Union, Washington, DC.

Gilbert, R.O. 1987. Statistical methods for environmental pollution monitoring. Van Nostrand Reinhold Co., New York.

Hannan, E.J. 1955. A test for singularities in Sydney rainfall. Austr. J. Phys. 8:289-297.

Hinkley, D.V. 1970. Inference about the change-point in a sequence of random variables. Biometrika 57:1-17.

Hipel, K.W. 1981. Geophysical model discrimination using the Akaike information criterion. IEEE Transaction on Automatic Control AC-26:358-378.

Hipel, K.W., W.C. Lennox, T.E. Unny, and A.I. McLeod. 1975. Intervention analysis in water resources. Water Resour. Res. 11:855-861.

Hipel K.W., A.I. McLeod, and E.A. McBean. 1977. Stochastic modeling of the effects of reservoir operation. J. Hydrol. 32:97-113.

Hirsch, R.M. 1979. Synthetic hydrology and water supply reliability. Water Resour. Res. 15:1603-1615.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. Techniques of trend analysis for monthly water quality data. Water Resour. Res. 18:107-121.

Jenkins, G.M., and D.G. Watts. 1969. Spectral analysis and its applications. Holden-Day Series, In Time Series Analysis, San Francisco.

Koch, R.W., and G.M. Smillie. 1986. Bias in hydrologic prediction using log-transformed regression models. Water Resour. Res. 22:717-724.

Lane, W.L. 1975. Extraction of information on inorganic water quality. Hydrology Papers, No. 73, Colorado State University, Fort Collins, CO.

Lee, A., and S.M. Heghinian. 1977. A shift of the mean level in a sequence of independent normal random variables -- A bayesian approach. Technometrics 19:509-506.

Lee, E.S. 1972. Analysis, modeling and forecasting of stochastic water quality systems, Vol. I, time series analysis in water quality modeling. Water Resources Research Institute, Kansas State University, NTIS PB-226566, 272 pp.

Lettenmaier, D.P. 1976. Detection of trends in water quality data from records with dependent observations. Water Resour. Res. 12:1037-1046.

Lettenmaier, D.P. 1977. Application of intervention analysis to environmental impact assessment, presented at Symposium on Methodologies for Environmental Assessments in Energy Development Regions.

Litwin, Y.J., and E.F. Joeres. 1976. Stochastic modeling of bi-hourly river dissolved oxygen records monitored at a fixed cross-section, In Mathematical Models for Environmental Problems. Proceedings of the Int. Conference held at the University of South Hampton, England, Sep. 8-12, 1975, C.A. Brebbia ed.

Loftis, J.C., and R.C. Ward. 1980a. Sampling frequency selection for regulatory water normality monitoring. Water Resour. Bull. 16:501-507.

Loftis, J.C., and R.C. Ward. 1980b. Water quality monitoring--some practical sampling frequency considerations. Environ. Manage. 4:521-526.

Loftis, J.C., R.H. Montgomery, J. Harris, and others. 1986. Monitoring strategies for ground water quality management. Colorado Water Resources Research Institute Completion Report No. 140, 125 pp.

Loftis, J.C., and R.C. Ward. 1987. Designing systems to monitor water quality impacts due to non-point sources. Water Management: Proceeding from the 1987 Regional Meeting Sponsored by the U.S. Committee on Irrigation and Drainage, Denver, CO., pp. 1-6.

Loucks, D.D., J.R. Stedinger, and D.A. Haith. 1981. Water resource systems planning and analysis. Prentice-Hall, NJ.

McDonald, G., J. Jones, G. Breazeale, and others. 1985. Water quality standards: Bull Run Watershed. Forest Service, U.S. Department of Agriculture.

McLeod, A.I., K.W. Hipel, and F. Camacho. 1983. Trend assessment of water quality time series. Water Resour. Bull. 19:537-547.

Matalas, N.C., and D.R. Dawdy. 1964. Handbook of applied hydrology, V.T. Chow, ed. Ch. 8, McGraw Hill, New York.

Mood, A.M., F.A. Graybill, and D.C. Boes. 1974. Introduction to the theory of statistics. McGraw-Hill, New York.

Parzen, E. 1967. Time Series Analysis Papers. Holden-Day, San Francisco.

Plate, E.J., and L. Duckstein. 1988. Stochastic aspects of water quality modeling for non-point sources, paper presented at the International Symposium on Water Quality Modeling of Agricultural Non-Point Sources, Utah State University, Logan, UT.

Salas, J.D., and D.C. Boes. 1980. Statistical computer techniques in hydrology, Unpublished Notes, Department of Civil Engineering, Colorado State University, Fort Collins, CO.

Salas, J.D., D.C. Boes, and R.A. Smith. 1982. Estimation of ARMA models with seasonal parameters. Water Resour. Res. 18:1006-1010.

Salas, J.D., J.W. Delleur, V. Yevjevich, and W.L. Lane. 1980. Applied modeling of hydrologic time series. Water Resources Publications, Littleton, CO.

Salas, J.D., and R.A. Smith. 1980. Correlation properties of periodic AR(p) models, Proceedings of the III International Symposium on Stochastic Hydraulics, Tokyo, Japan, pp. 107-115.

Sanders, T.G., and D.D. Adrian. 1978. Sampling frequency for river quality monitoring. Water Resour. Res. 14:569-576.

Sanders, T.G., R.C. Ward, J.C. Loftis, and others. 1983. Design of networks for monitoring water quality. Water Resources Publications, Littleton, CO.

Sen, A., and M.S. Srivastava. 1975a. Some one-sided tests for change in level. Technometrics 17:61-64.

Sen, A., and M.S. Srivastave. 1975b. On tests for detecting change in mean. Annals of Statistics 3:98-108.

Snedecor, G.W., and W.G. Cochran. 1980. Statistical Methods. The Iowa State University Press, 507 pp.

Tabios, G. 1984. Non-linear stochastic modeling of river dissolved-oxygen, Ph.D. dissertation, Colorado State University, Fort Collins, CO.

Tao, P.C., and J.W. Delleur. 1976. Seasonal and non-seasonal ARMA models in hydrology. J. Hydr. Div. ASCE 102:1591-1559.

Thomas, H.A., and M.B. Fiering. 1962. Mathematical synthesis of streamflow sequences for analysis of river basin by simulation. In A. Mass, et al., eds, The Design of Water Resources System, p. 459-493 Harvard University Press, Cambridge, Mass.

Wastler, T.A. 1963. Application of spectral analysis to stream and estuary field surveys 1. Individual power spectra, Publication N. 999-Wp-7, U.S. Public Health Service, Washington, DC.

Yevjevich, V. 1972a. Structural analysis of hydrologic time series. Hydrology Paper 56, Colorado State University, Fort Collins, CO, 59 pp.

Yevjevich, V. 1972b. Stochastic processes in hydrology. Water Resources Publications, Littleton, CO.

Yevjevich, V. 1972c. Probability and statistics in hydrology. Water Resources Publications, Littleton, CO.

STOCHASTIC ASPECTS OF WATER QUALITY MODELING FOR NONPOINT SOURCES

Erich J. Plate[1], and Lucien Duckstein[2]

ABSTRACT

Practical applications of water quality models require to take the stochastic variability of the model components into account. In this paper, types of water quality models are classified according to their structure, and sources of stochasticity are identified. We shall distinguish stochasticity due to natural variability, which is caused by the randomness of a natural process, and uncertainty, which is a result of lack of information on the true state of nature. Water quality models are based on hydrologic input models of varying complexity, which are converted to water quality models by means of process functions describing the relation between water quantity and quality transport. Simple examples are given. It is emphasized that the importance of the influence of the stochasticity must be evaluated relative to the purpose of the model.

INTRODUCTION

Water resources research of the last decades has become almost synonymous with computer modeling. This applies in particular to research on water quality problems resulting from nonpoint-source pollution. Most of this research has concentrated on deterministic models, and models of great complexity have been built to describe pollution situations for clearly defined initial and boundary conditions. However, in the real world there are no true deterministically defined pollution situations; almost everywhere stochastic effects intrude. In fact, it is possible that for many water quality problems the statistical uncertainty is so large that it invalidates the benefits from refined modeling. In such cases, it makes sense to obtain a trade-off of deterministic against stochastic model improvement. This requires in the first place that stochastic aspects be incorporated and considered in model design and development.

In recent years, much progress has been made in identifying and defining sources of stochasticity in hydrologic and in water quality models. An extensive recent review of the analysis of uncertainty in water quality modeling has been given by Beck (1987), who presented a taxonomy of uncertainty complete with elaborate discussions of the methods which can be used to handle uncertainties. The purpose of this paper is less ambitious: we intend to discuss some of the usual concepts of uncertainty and variability and to show, by means of a few general examples, how to obtain quantitative information in special cases of water quality models (WQMs). Because deterministic components of WQMs, such as the submodels describing the transfer relations between water movement and pollutant transport, have been covered extensively in other contributions to this Symposium, we shall concentrate on stochastic aspects.

Nonpoint-source pollution problems considered relate to hydrologic basins. A typical hydrologic basin is shown in figure 1. It has subareas differing in plant cover, soil texture, or topography. Water movement forms the agent for the transport of most substances in this area. It is described by a hydrologic model. To obtain a WQM, this hydrologic flow model is combined with a model of pollutant emissions for calculating the pollutant transport. Pollutants may be one or more of many substances, ranging from air pollution deposits to natural sediments (ASCE Task Committee, see Kelman et al. 1977). The pollutant is assumed to emanate from an areal source of

[1]Erich J. Plate, University of Karlsruhe, Karlsruhe, FRG.

[2]Lucien Duckstein, Case Western University, Cleveland, Ohio, USA.

pollution which exists in one or more of the subareas. Its movement from the area of origin to other points within the basin is described analytically by means of WQMs. The purpose of a WQM is the prediction of the time distribution or some of its parameters (such as the mean value) of the concentration of pollutants at one or more points on or below the earth's surface, or in the runoff from the surface or the exfiltration from the groundwater.

Model Components of WQMs

For the purpose of our study it is important to distinguish the elements of a WQM sketched in figure 2. The dynamic model input consists of the water input whose origin for a hydrologic basin is precipitation, and the pollutant input. The input or inputs are transformed by process models into modified functions of space and time: for example, the soil converts rainwater and rainwater pollution, such as acidity, into a recharge of the groundwater of different acid concentration; or the basin as a hydrologic filter transforms the rainfall into runoff at the basin outlet. Process models quantify the causal physical chain of events from rainfall and generation of pollutants to concentrations of polluting substances in surface or groundwater.

Another important aspect is shown in Figure 2: the decision model, which should be considered as part of a WQM. WQMs usually are constructed for certain applications, i.e. they are developed as decision tools. Their objective is to aid decisions aimed at preventing pollution or at lowering contaminant concentrations, or to investigate the effects of policy decisions (Milon 1987).

Stochasticity in WQMs

Stochasticity of WQMs exists basically in two forms: the stochasticity due to natural variability (where natural variables are realizations of a stochastic process) and the stochasticity due to uncertainty (which results from insufficiency of information) despite the fact that this information could be obtained. The big difference between the two types of stochasticity is that we have to live with the natural variability and must incorporate its effect into the decision processes: our decisions are a gamble against nature, or of risk taking. Uncertainty, on the other hand, is subject to our own choice - at least theoretically.

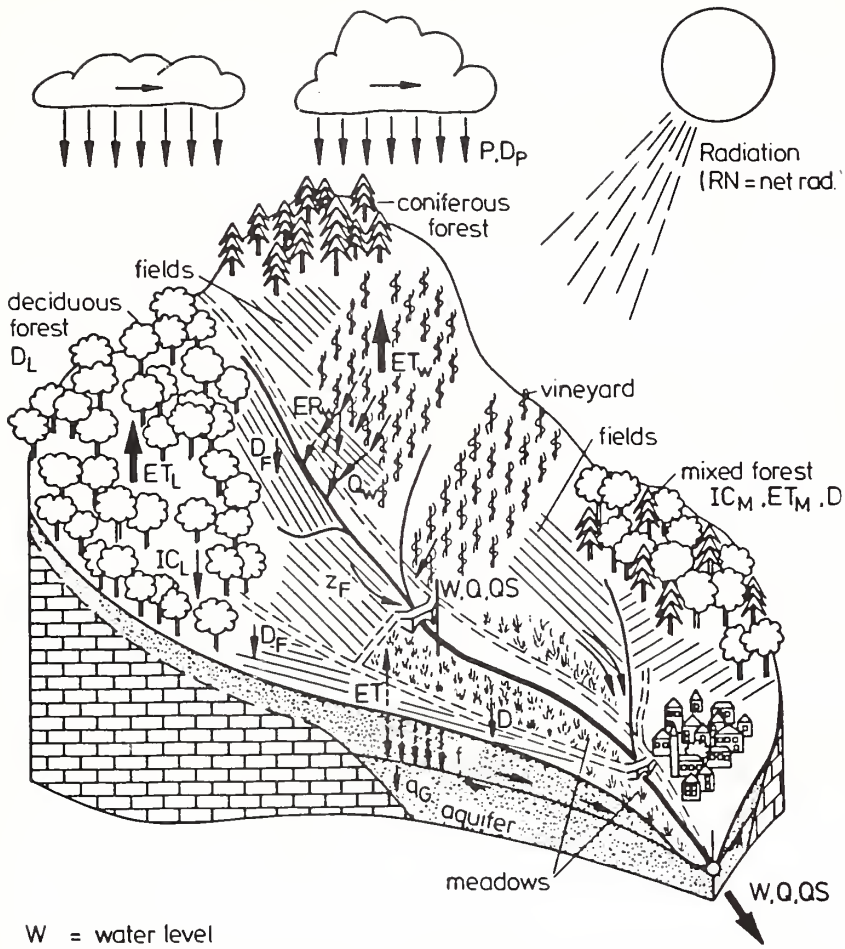
Natural Variability of WQMs

The most important source of stochasticity is the natural variability affecting components of a WQM. For example, external inputs such as storm rainfall intensity, exhibit a natural variability in space and time. This variability is always present, and no matter how accurately a process with natural variability is measured, it is not possible to get rid of its stochasticity. The best description obtainable is an analytical model for the random process of which the actual situation is a realization.

Uncertainty in WQMs

Uncertainty, in the terminology of this paper, is defined as the stochasticity associated with our ignorance. The implication is that it could be avoided - if, for example, a very large amount of effort and time could be spent on improving measuring instruments, increasing the number and duration of measurement time series, and of improving the accuracy of models and of calculations. Uncertainty may stem from different causes: it is customary to distinguish between data uncertainty, model uncertainty, and parameter uncertainty.

Effect of data uncertainty. A primary source of stochastic uncertainty is the inadequate quality of the data base. It occurs in three forms: as sample uncertainty, which results from the fact that not enough input data are available to infer exactly the ensemble characteristics of the random variables which are observed; as measurement uncertainty which results from the inability to measure exactly the value of individual data points; and as resolution uncertainty which is caused by temporally or spatially variable quantities being measured only at discrete points.



W = water level
P = precipitation
ET = evapotranspiration
IC = interception
f = infiltration
z = erosion from field
q = runoff from surface
Q = discharge of creek
D = vector of pollutant inputs (D_P from precipitation)
QS = vector of pollutant transport
ER = erosion yield

Index P = precipitation
F = field
W = vineyard
L = deciduous forest
M = mixed forest
N = coniferous forest
G = ground water

Figure 1.
Schematic representation of a hydrologic basin with nonpoint sources of pollutant (QS consists of: ER=erosion yield, FE=fertilizer effluent, PE=pesticide effluent, GM=geomorphologic substances).

Effect of model uncertainty. A second source of stochasticity lies in the uncertainty of the model. It is not possible to represent exactly a situation such as the one shown in figure 1 by a physical model. WQMs must be simplified in order to be tractable. This may lead to an acceptable representation of the input-output relations considered for the case of calibration, i.e. for the fitting of the model to an observed set of data. It is often found that simplified models also yield acceptable validations, when validation calculations are done under similar conditions as had been used for calibration. But this does not justify the assumption that a simplified model can be extrapolated to extreme cases, because processes which have been masked completely during

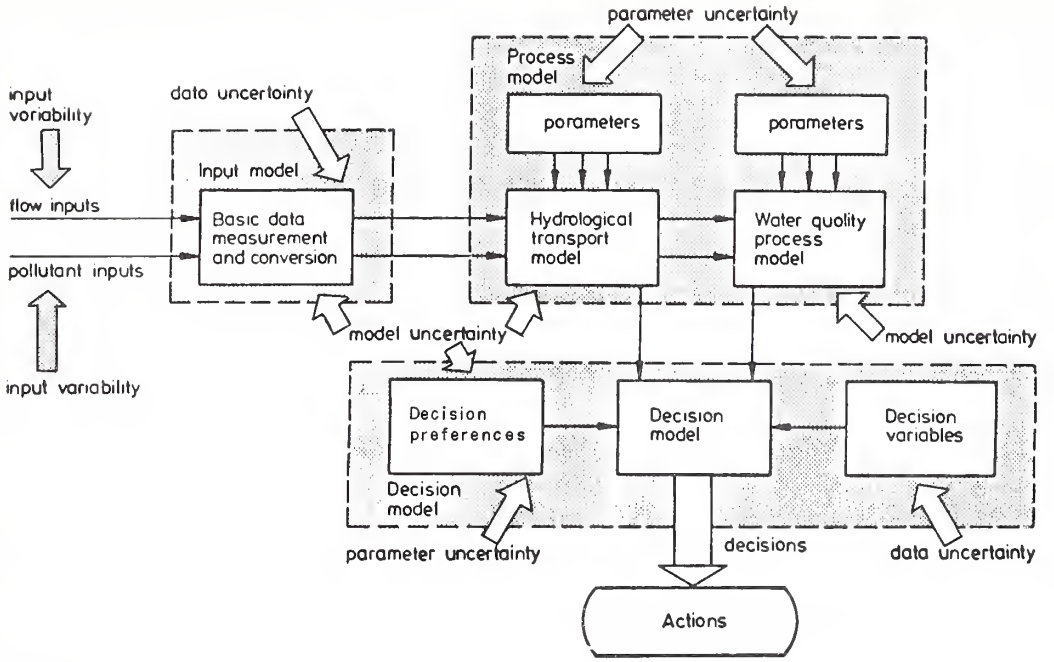


Figure 2.
Schematic representation of a water quality model and associated sources of uncertainty.

calibration conditions may dominate during decision conditions (Sorooshian and Gupta 1983). Consequently, models which have to be extrapolated (or transferred from one area to another, or from the laboratory to the field), induce uncertainty into WQMs. A part of the model uncertainty is also the numerical uncertainty which arises because discrete time steps and finite grid sizes must be used in numerical modeling. This model uncertainty shall not be discussed here.

Effect of parameter uncertainty. A simplified model for operational purposes generally requires empirical parameters with no or only a small physical basis, which must be obtained from calibrations. For some models (which are called conceptual models, see below), the parameters are determined from measured input and output functions. Such model parameters are random variables, usually with unknown probability distribution functions. Calibrations of such models by means of measurements usually are done to find the best constants for model parameters (Troutman 1985), and any natural parameter variability that might be present is attributed to model or data uncertainty.

From these introductory remarks it becomes evident that we must be concerned with the natural variability of the input, as the first order source of stochasticity. The hydrologic system responds to the random input in a basically deterministic fashion, through a chain of process models. Both the inputs and the process models are subject to the types of uncertainty identified above, as will be discussed further in the following sections.

INPUT MODELS FOR WATER QUALITY MODELING

The external input (see figure 2) has two parts: the input through rainfall and other meteorologic variables (which provides the energy for the transport mechanism) and the input or source function of pollutants, i.e. the externally induced space-time distributions of pollutants by quantity. These are the most important sources of natural variability, and of primary concern in the consideration of stochasticity of WQMs.

The external WQM input can be either a point or a distributed input. Point inputs typically consist of a vector of input functions of time. The process model converts this input function vector into a vector of output functions. The transformation through the process model depends on the structure of the input function.

Natural Variability of Inputs

In a natural environment, the WQM input is usually a stochastic process, whose realization consists of random functions in space and time. An important characteristic of input random processes is that they are well defined (by their realizations, although their physical structure is perhaps unknown). We distinguish three important input types, which are shown schematically in figure 3.

Probabilistic Inputs.
Figure 3a shows the case of a sequence of n single valued independent inputs, x_i, which are not correlated. They are usually taken one each from one time interval ΔT, and are called probabilistic inputs. Typical examples for this type of input are the sequences of extreme annual floods, droughts, or extreme values of rainfall magnitudes, which serve, usually with $\Delta T = 1$ year, as inputs to many design models. The probabilistic description of such inputs is expressed by a single function, the probability density function (pdf) of the magnitude. The use of different extreme value distributions to describe this pdf is standard hydrologic procedure (Chow 1964, Haan 1977), application to elementary process models (which are defined below) is standard engineering practice (for example, for the design of the spillway of a dam).

Event Inputs.
Figure 3b shows an input which consists of a sequence of events i(t) which are separated by time intervals τ_r of no activity. This is the case of an event input. Statistical models for event-based inputs and deterministic process functions have a rather old tradition in hydrologic applications, where usually single events, such as extreme rainstorms, have been considered as inputs to rainfall-runoff models of unit hydrograph type. The stochasticity of the input is expressed by assuming an average shape for the event distribution over time, and assigning a probability of occurrence to the event - either according to peak value, or according to depth and duration.

For water quality models, one usually encounters process models which are nonlinear and have uncertain characteristics. For inputs to such process models it is required to use the pdf of the event. A typical example for this type of inputs is the sequence of hyetographs which are generated by convective (thunderstorm type) rainstorms (Duckstein et al. 1972). One approach to the problem is to assume that the input is a rectangular effective rainfall pulse $P_r=i_r t_r$ of intensity (pulse height) $i_r(t)$ and duration t_r, where both P_r (or i_r) and t_r are random variables, as indicated schematically in figure 3b by the dashed rectangles for the events. Many authors (Cordova and Bras 1981, Cordova and Rodriquez-Iturbe 1985, Julien and Frenette 1987) assume these random variables to be uncorrelated and exponentially distributed. Other models have also been used, such as the one of Woolhiser and Osborn (1985), who used a random function to describe the distribution of the intensity i(t) over time.

Note that the case of an event input is a special case of a more general input process in which two (or more) processes alternate at random intervals of time. Such processes are called intermittent and find important applications in fluid mechanics and hydrology (Plate 1976, Yevjevich 1984). An important contribution to models of this type has been the model of Todorovic and Zelenhasic (1970). This model adds a random variable for the time between events, and has been the starting point for many later models of on-off type processes.

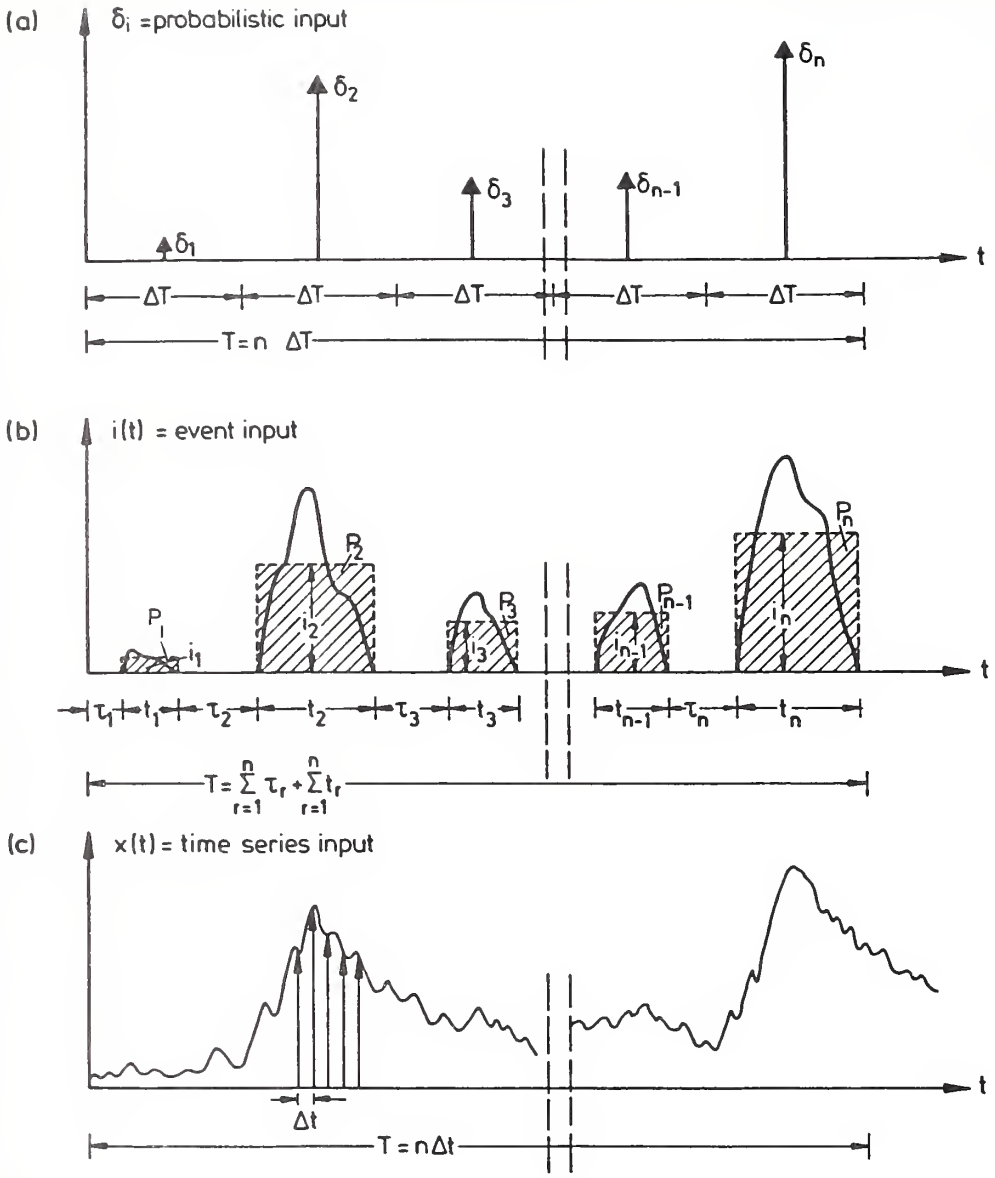


Figure 3.
Types of input functions for WQMs.

Time Series Inputs.
Figure 3c represents the case of an input consisting of a continuous time function, called time series input. Examples of this type of process are runoff hydrographs for rivers, and many other random geophysical processes (Yevjevich 1984). Time series models are seldom used for short-term descriptions of hydrologic inputs. However, they find application for long-range models as used, for example, for the design of water supply reservoirs. Although only few applications of this type of input to WQMs are known to the authors, there is no question that the generation of long-term time series may become important in the future for simulating long-term effects of water pollution (see for example Rosso 1986, for a discussion of the role of time series analysis in sedimentation problems).

Input Variability for Field Processes

The most complex case is the representation of the input by a space-time variable stochastic field. One of the most active areas of research is the modeling of rainfall fields. For a survey of the

literature, reference is made to Waymire and Gupta (1981), and to Plate (1986). It may suffice to state that such models could be useful for stochastic pollution simulation by accounting for spatial as well as for temporal variations of input functions. If the continuous field of some input is known, areal averages can be determined by integration and thus the uncertainty caused by ignoring the spatial variability can be quantified. However, such knowledge is not likely to have a large effect on decisions: in essence, all such an analysis may accomplish is to shift part of the stochasticity from uncertainty to natural variability.

PROCESS MODELS IN WATER QUALITY MODELING

In general, process models describe the relationship between physical quantities in terms of an analytical expression, or they describe chemical and biological relations between variables. Both types of models can be deterministic, i.e. they are set up under the assumption that the models in all their aspects are either exactly known, or can be approximated by a realistically simplified model whose parameters can be determined once and for all by a calibration. Or else, they are stochastic, in which case part or all the uncertainty and natural variability of a natural process is accounted for through statistical descriptions.

The difference between a deterministic model of any kind, and its stochastic equivalent is simply that there exists a one-to-one correspondence between input and output for a deterministic model, whereas a stochastic model converts any type of input into random output variables. Therefore, the quantities to be considered are not the values of the outputs themselves but functional form and parameters of their probability density distributions.

The analytical or numerical treatment of the stochastic aspects of a WQM depends very much on the type of process model, and on the type of input function which is used. Furthermore, it will depend on the type of output which is required. The choice of the model will have to depend on the objective of the study, but in general it can be said that the best model is the one that gives the desired result with sufficient accuracy at minimum effort in cost and time. In the following section, stochastic aspects of some typical combinations of input functions and process models are discussed.

Elementary Water Quality Process Models

The simplest type of process model consists of a deterministic one-to-one relationship between input variable (x) and output variable (y), i.e., a relation of the form:

$$y = g(x). \qquad [1]$$

Such a relation is assumed to exist for many water quality processes. The most frequently used elementary relations for process models of pollutants are of the form:

$$y = Ax^B \qquad [2]$$

where A and B are (often empirical) parameters. The most elementary case is the linear transformation:

$$y = ax + b. \qquad [3]$$

This equation is a satisfactory analytical WQM for many (conservative) pollutants. For example, in the case of pure dissolved pollutants, x is the discharge or the rainfall, and y is the output concentration (Haith and Tubbs 1981). As another example, for the case of a pollutant which

adheres to the soil moved by surface erosion - such as phosphorus (Duckstein et al. 1978), one finds that x is the eroded soil and y is the outflow of phosphorus.

Natural Variability of Elementary Water Quality Process Models

The linear process model with random input is, mathematically speaking, a simple one-to-one transformation of a probabilistic input x_i with probability density function (pdf) f(x) into a probabilistic output y_i with pdf f(y). The function f(y) or its parameters have to be determined by means of standard methods from probability theory. The process equation, equation 3, is a transformation of the input variable x (which could be the realization of a probabilistic point process) into the output variable y, with a resulting output pdf f(y) = f(x)dx/dy:

$$f(y) = (1/a)f(x) \quad \text{with } x = (y-b)/a. \qquad [4]$$

This transformation is represented graphically by a straight line, so the shape of the output pdf is the same as that of the input, and the moments of the output pdf are calculated from those of the input pdf by a linear transformation. For example, an input x with mean μ_x and variance σ_x^2 is transformed into an output y with mean $\mu_y=a\mu_x-b$ and variance $\sigma_y^2=a^2\sigma_x^2$, and a normal pdf for x converts to a normal pdf for y.

Linear relationships are also found frequently by a linearization of equation 2 which is obtained by taking the logarithms of both sides. For example, in sediment studies, this type of transformation is used to describe the relationship between the logarithms of river discharge (lnQ) and of its suspended sediment load ($\ln Q_s$) through a linear relation equivalent to equation 2:

$$\ln Q_s = A + B(\ln Q) \qquad [5]$$

with mean value $\mu_{\ln Q_s}=A+B\cdot\mu_{\ln Q}$ and variance $\sigma^2_{\ln Q_s}=B^2\sigma^2_{\ln Q}$, where B ranges from 1.5 to 3. An example of this kind of relationship is shown in figure 4, which has been presented by Walling (1977). The advantage of equation 5 is that the best-fitting straight line can be found from linear regression analysis. This analysis is based on minimizing the square of the deviations of the logarithms of the measured y values for given x from the values of y calculated from equation 5, i.e. it is based on deviations which are considered, on the average, to be proportional to the calculated value of y. This is well suited for data which have the property that the deviations from a regression curve increase approximately proportionally to the magnitude of the measured quantity. Furthermore, if Q is a lognormal variable, as is frequently found, then the transformation equation 5 produces lognormal variable Q_s.

Uncertainty of Elementary Water-Quality Process Models

The example of figure 4 illustrates the effect of uncertainty in elementary process models. It is typical of many linear relationships which are postulated to exist between two geophysical variables. Very often there exists no physically or otherwise well-founded reason for the linear relation. It cannot be discerned, in the absence of a correct physical model, if the deviation of the individual data point from the regression curve is caused by measurement error, or by not properly accounting for the true process. The true process may be hidden because it may depend on other parameters, which have not been recognized to have an effect, or because the curve relating the two variables is not linear (in the original plane formed by the two measured quantities, or in a plane to which the quantities have been transformed). All these might explain the scatter of data points around the average curve, and thus cause the uncertainty of the process model, which could be overcome by better models or measurements.

In the case of equation 5 the effects of all parameters describing basin and soil characteristics as well as the erosion and transport processes are accumulated into A and B. However, it is evident, from the large scatter of the data, that effects of other variables exist. Since they cannot be identified specifically, they contribute to the non-uniqueness of the input-output relationship.

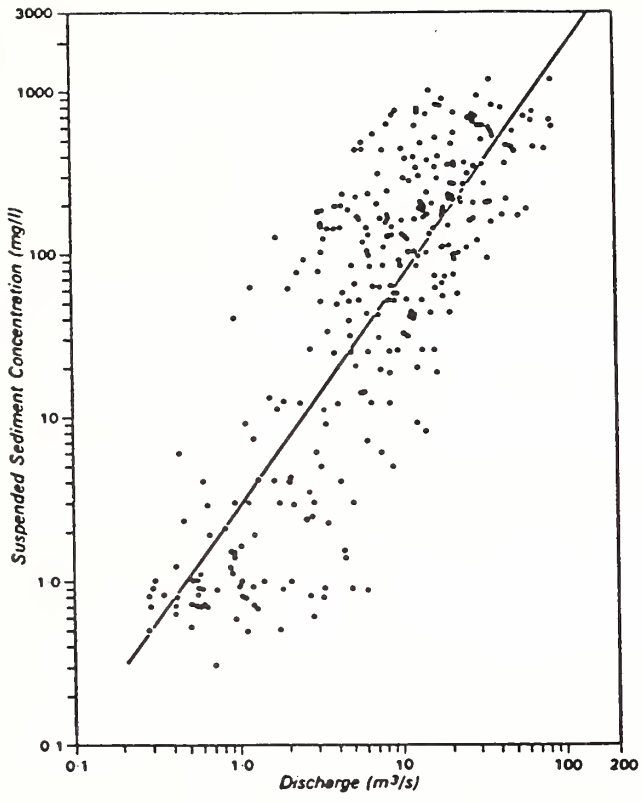


Figure 4.
Example of a linear regression result of the logarithms of suspended sediment load and discharge (from Walling 1977).

This is an example which shows that model refinements can be traded off against uncertainty, i.e. the deterministic relationship in equation 5 becomes a stochastic transformation.

In the case considered uncertainty can be handled in two ways. First, one can assume the scatter of the data to be associated with an empirically determined random process ϵ, with mean value 0 and variance σ^2_ϵ. Then equation 5 becomes:

$$\ln Q_s = A + B \ln Q + \epsilon \qquad [6]$$

If ϵ is assumed to be a normal random variable with zero mean, then Q_s remains a lognormally distributed variable with mean $\mu_{\ln Q_s}$. Its variance, however, is increased to:

$$\sigma^2 = \sigma^2_{\ln Q_s} + \sigma^2_\epsilon \qquad [7]$$

As a second alternative, if more is known about the structure of the process model - for example if other parameters are known to influence the transformation - then A and B may be assumed random variables A′ and B′ which obey a pdf f(A′) or f(B′). The expected values of A′ and B′ are usually taken as the linear regression estimates. Assume for simplicity that A is a true constant, and only B is a random variable. Then the pdf of Q_s is found from the formula:

$$f(\ln Q_s) = \int_{-\infty}^{\infty} f(\ln Q_s | B)\, f(B)\, dB \qquad [8]$$

where $f(\ln Q_s | B)$ is the conditional pdf for $\ln Q_s$ given B. The solution to this equation depends on the shape of f(lnQ) and f(B) and in general has to be determined by numerical integration.

Extending the regression model, one may use empirical relationships between the principal variables as determined from experiments. The conversion of the pdf through a nonlinear transformation changes the shape of the output pdf, and usually the calculation of the output pdf is a complicated procedure, which has to be done numerically or by simulation. This also holds for dynamic variability, in which case the conversion takes place through stochastic differential equations, which only recently have been considered in the context of hydrology.

Hydrologic Models as Basis for WQMs

Process models relate an input variable to an output variable. Most WQMs have process components in which an output variable, y, describing the transport of pollutant is related to the variable x, describing the flow of transporting agent. The transporting agent usually is water, which transports pollutants directly, as in the case of dissolved salts, or indirectly, by transporting the sediment which in turn transports the substances adhering to it. The basis of a WQM is therefore the quantification of the various aspects of the rainfall runoff process, and all associated stochastic modeling questions also pertain to WQMs. It is useful to classify WQMs according to the hydrologic model which is used as a transport model for the water quality variables.

The input-output model which converts rainfall into runoff is a typical process model. We distinguish on the one hand between area element models, homogeneous area models, and field models, which are models described by physical process equations, and conceptual models on the other hand, which are empirical. The water quality aspects add, for the substances considered, transfer functions to the model of the water motion, and additional continuity conditions imposed by the total available pollutant.

Water Quality Models of Area Elements

The smallest unit of a WQM is an element of the basin, as indicated schematically in figure 5, in which the mass transport of water due to rainfall is determined by the infiltration characteristics. The transport-effective fraction of the total rainfall is divided into a vertical infiltration component, and (after ponding) a horizontal surface runoff component. The infiltration replenishes the groundwater and carries surface-deposited pollutants downward and, after some filtering through the unsaturated layer of the soil, results in groundwater pollution. The surface runoff carries the pollutant to neighboring elements under the effect of gravity; the effect of lateral transport through the unsaturated zone is usually not considered. Many models exist to describe the rainfall-runoff process for an area element, such as the model presented by Maniak (1986).

Stochastic models to describe the process of pollutant infiltration exist. An extensive discussion of the statistical analysis of soil moisture transport in soils including the transport of conservative pollutants has been given in a series of papers by Dagan and Bresler (1983; see also Boulier and Vauclin 1984). The general solution for soil moisture transport is very difficult to obtain: usually the analysis is restricted to simple initial and boundary conditions, such as the condition of constant infiltration rate, or constant moisture content at the surface, and the general case of stochastic variability of the moisture content as function of random input processes, such as rainfall has not been treated. For the cases investigated, it is interesting to note that it is possible to describe the stochastic variability of the profile parameters for moisture, soil tension, and conservative pollutant through a single variable. As Dagan and Bresler have found out from field experiments with natural soils, for a homogeneous area this variable is lognormally distributed.

Water Quality Models of Hillslopes

The element from which more general physically based hydrologic models must be generated are hillslope models, in which a slice is taken out of a part of the terrain along the gradient of the surface. We distinguish between earlier, agricultural models, which we shall call "homogeneous area models", in which the hillslope was an agricultural field, and recent models in which hillslopes of more general type are considered.

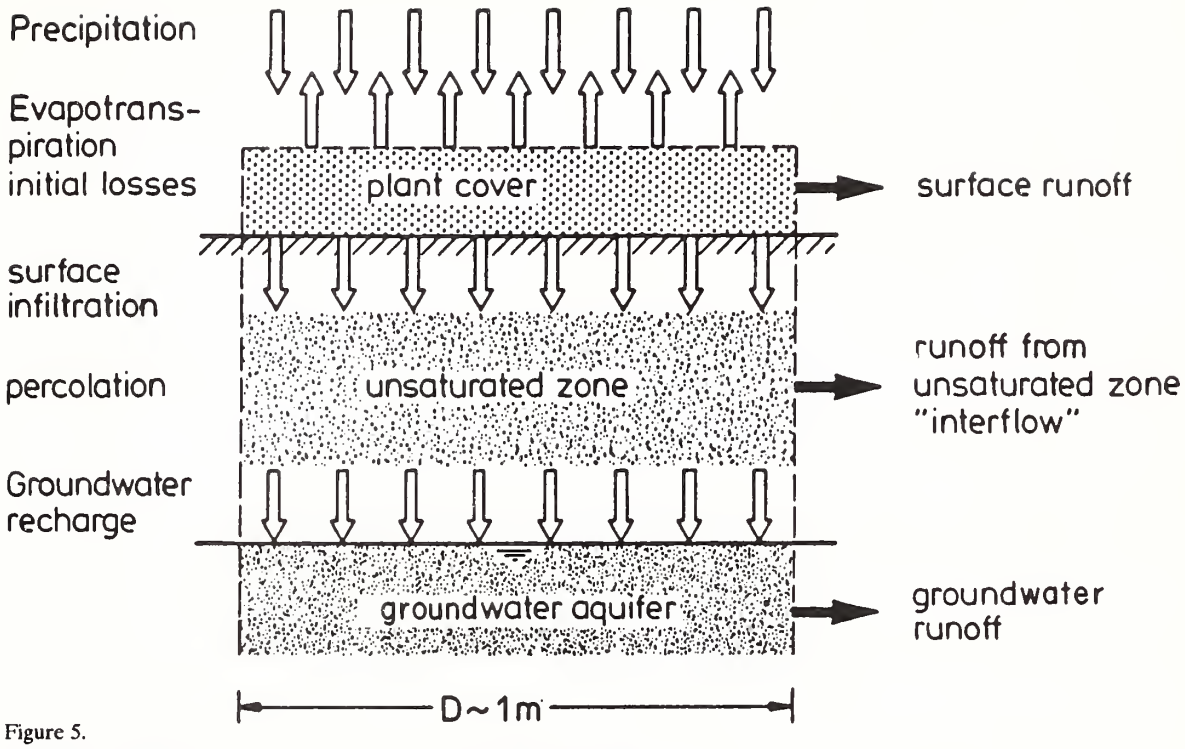


Figure 5.
Schematic representation of a WQM based on a hydrologic-area element.

Homogeneous area models. An important special case of a catchment is an area of homogeneous soil composition and of constant slope with a single type of crop covering the whole field, as indicated schematically in figure 6. The runoff from such an area can be approximated by a hydraulic relation, and a function similar to equation 4 is often used, in which parameter a depends on the geometry and the roughness of the field.

Most of the research on pollution from agricultural sources has been done on homogeneous areas, in particular the erosion work, which resulted in the Universal Soil Loss Equation (USLE). This well-known equation in its original form (or in the modified form given in Williams (1975), which is the modified USLE, or MUSLE) forms the basis of much work on stochastic sediment research (Smith et al. 1977, Bogardi et al. 1985). For example, Smith et al. (1977) have combined the MUSLE with the SCS method of calculating rainfall abstractions to obtain the following expression for the sediment yield z(r) per rainfall event r:

$$z(t) = W \left[\frac{a_o P_r^4}{(P_r+f)^2(a_1 t_r + t_c)} \right]^{0.56} \qquad [9]$$

where w and a_o = conversion factors incorporating size of field, land use, etc.
P_r = effective rainfall depth per event r [mm] for r=1,2...,N where N is number of events per season
f = constant infiltration [mm]
a_1 = dimensionless constant
t_r = rainfall event duration [hrs], and
t_c = time to concentration.

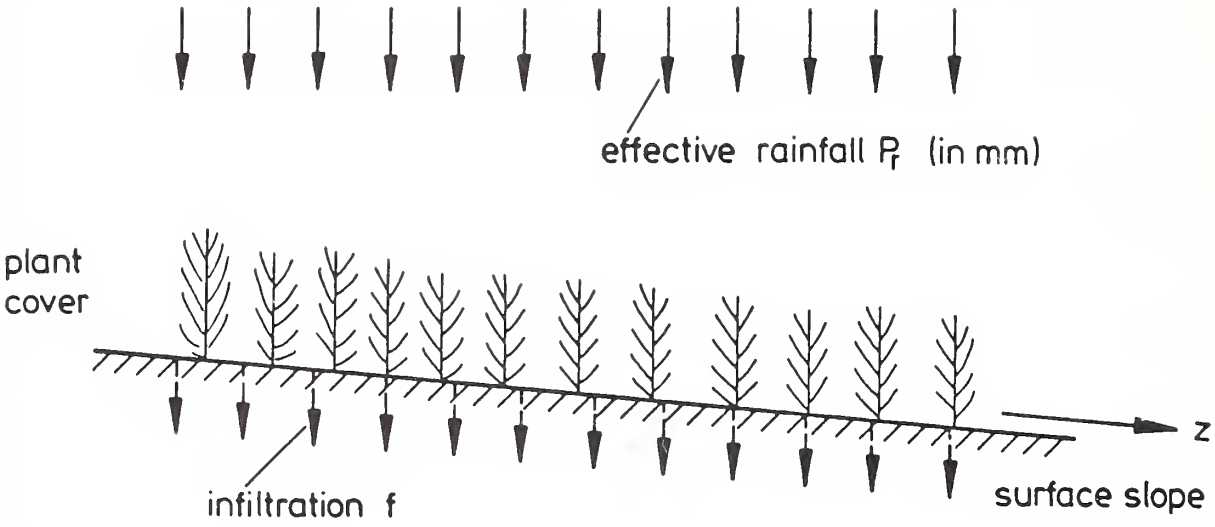


Figure 6.
Simplified representation of a homogeneous hydrologic area.

For details and for numerical values, the original paper by Smith et al. (1977) or Bogardi et al. (1985) should be consulted.

The structure of equation 9 makes it clear that the input for this model must be an event-based rainfall, as shown in figure 3b. It has been solved numerically (Smith et al. 1977) with rainfall event inputs which have a joint pdf for rainfall intensity and rainfall duration described by a bivariate exponential pdf $f(P_r,t_r)$, with Poisson-distributed number of events (N) per season. It has also been solved by simulation in Bogardi et al. (1985). Both groups of authors also considered the uncertainty of the parameters which are used both in the input as well as in the process models. The method of solution is schematically illustrated in figure 7. The random variables P_r and t_r are the coordinates of a plane in which isolines of the joint pdf $f(P_r, t_r)$ are indicated. In this plane, the condition z=constant calculated from equation 9 leads for different values of the constant to a family of curves shown schematically in figure 7. The cumulative distribution F(z) is calculated for each z=Z by integration of $f(P_r, t_r)$ over the area in which z<Z, as indicated by the unshaded region of figure 7. Finally, the probability density function f(z) is found from F(z) by differentiation.

Note that water quality parameters which are directly proportional to the rate of surface erosion are obtained by combining the soil loss equations with an elementary water quality model. The example of the phosphorus load of Bogardi et al. (1985) has already been mentioned. However, uncertainty is introduced by the fact that such a model is based on the assumption that the pollutant is distributed uniformly throughout the top soil layer, so that the amount of pollutant removed from the field by event r is equal to p z(r).

A more physically-based model for the surface erosion process has been employed by Julien (Julien and Frenette 1985, Julien and Dawod 1987), who used a more elaborate input hydrograph and an overland erosion model based on hillslope hydraulics.

The extension of area-element models to homogeneous areas (such as for agricultural fields) is of foremost interest also for WQMs. Fortunately, as Bresler and Dagan (1983) have pointed out, most applications do not require detailed knowledge of the actual spatial distribution of water

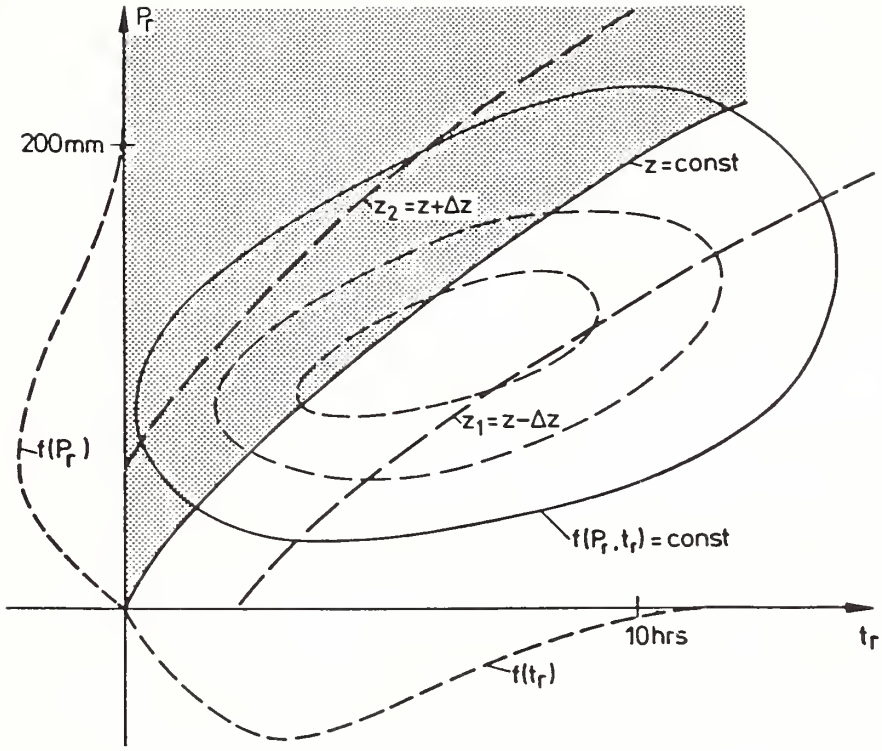


Figure 7.
Schematic representation of dependency of F(z) on the bivariate density distribution $f(P_r, t_r)$.

quality parameters; usually only the first, and occasionally the second moment of concentrations at significant points of the basin are needed.

Hillslope models. For many pollutants, the assumption of pollutant transport proportional to the quantity of surface erosion or water transport is not justified. It is necessary to determine the path of the pollutant transport, both to evaluate the filtering which the pollutant encounters on its path, and the residence time of the pollutant in soil or groundwater, or in the sediment deposits of a river. Models which permit such evaluations have to be more detailed than the homogeneous-area models, as is shown in figure 8. In figure 8, f is the infiltration rate at the surface, and f_n the recharge of the perched water above the groundwater table. The typical size of such a field is from a fraction of a hectare to a few hectares. This is the size for which detailed models of the rainfall-runoff process with constant parameters are particularly useful. As an example, the water transport model of Smith and Hebbert (1983) may be cited, in which surface and subsurface water movements are interacting through an infiltration equation, thus eliminating the need for an arbitrary specification of a runoff coefficient.

Hillslope models are not only more complicated from the physical point of view, they also are subject to many more uncertainties than the elementary area models of figure 5. All the parameters which describe the water transport in and above the soil vary randomly both in space and time. The parameters of a soil-moisture model depend on surface properties, such as plant cover and topography, or on soil properties, such as porosity or conductivity. They vary not only in space, but also to some extent in time, because of seasonal variations of parameters, which might also be affected by biological or chemical processes. The exact random process for the variability of internal parameters is usually not known, and has to be inferred from measurements. Because detailed measurements are usually possible only at a limited number of points in the field, the problem arises of extrapolating physical parameters from point measurements to area

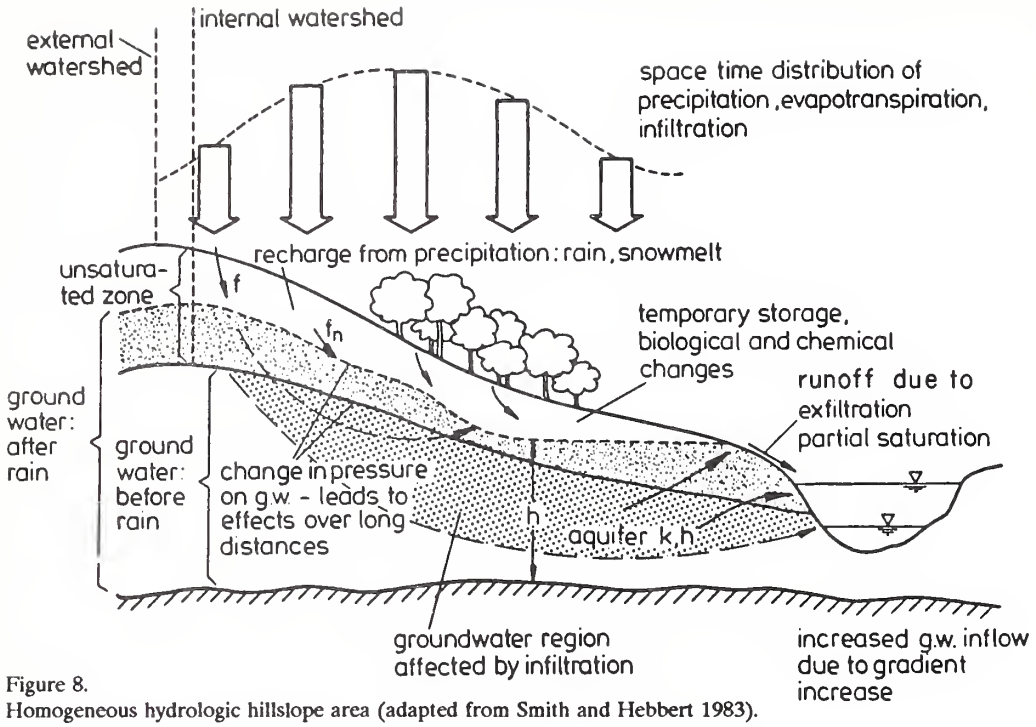


Figure 8.
Homogeneous hydrologic hillslope area (adapted from Smith and Hebbert 1983).

parameters. This is one of the most active areas of research in hydrology today, in which geostatistical methods such as kriging are extensively used.

Other sources of uncertainties for a homogeneous-area model of the strip type are illustrated schematically in figure 8. It is evident that the pollution from areal sources arriving in a creek cannot be correlated with the present time and areal distribution pattern of the input precipitation. It reflects a spatial pattern of its own which has developed over a long time span. For illustration, one notes that in figure 8 the flow in the creek before the rainfall event is fed directly by the groundwater through exfiltration. Therefore, the concentration in the creek is the average of the groundwater along the creek in the neighborhood of the shore. After a heavy rain the situation may have changed to yield a groundwater table as indicated by the short-dashed curve. In this case the pollutant which reaches the creek has many different sources. For example, near the creek two sources can be identified. The first results from the dry deposition which has entered the soil and which is washed out to the surface by exfiltration. The second originates from distant changes in the level of the water table: high infiltration rates in distant parts of the catchment cause increased pressure on the groundwater table. This pressure affects the entire aquifer and causes an increase in the groundwater discharge into the river, and consequently also of the pollutant which may be present in the groundwater. In the part of the catchment with exfiltration, the soil may be cleaned by the groundwater stream which passes through it. Wherever water infiltrates (for example in the upper part of the catchment) the pollutant which is deposited at the surface is washed into the soil and may in part be retained, so that the soil in this part of the catchment becomes more polluted.

Such flow situations have been observed, as is shown in the example of figure 9, for a non-ponding infiltration of acid deposition in Swedish till soil (Jacks et al. 1984). Before snowmelt, or fall rains, the discharge into the creek has a high pH value close to neutral, which results from the effluent from low-level groundwater. In the upper soil zone, the soil is acidic. Due to tillage, it has a hydraulic conductivity which decreases rapidly with depth. Therefore, during snowmelt or heavy rains, the flow into the creek results mostly from the upper soil layers and mixes with groundwater to yield a strong decrease in pH-value of the water - in spite of the larger flow, the

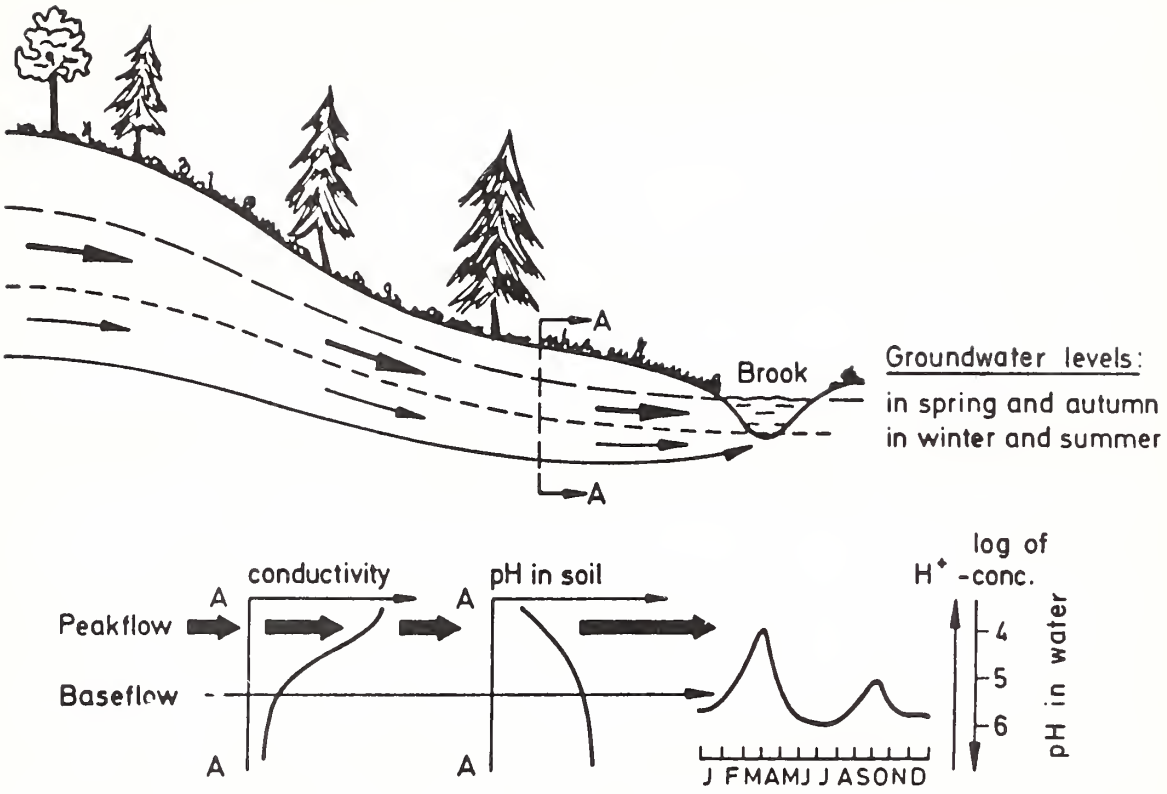


Figure 9.
Seasonal hydrogeology and chemistry of a tilled slope (from Jacks et al. 1984)

acid concentration increases! It is evident that the tracing of the pollutant complicates the model, because pollutants follow the particles of the flow field, and not the discharge, requiring their path to be calculated from a Lagrangian flow model.

In a typical terrain with random changes in elevation, a reversal of infiltration and exfiltration may occur not only in time but also in space and will cause a spatial pollutant distribution as a function of time which cannot with any reasonable effort be traced by an operational model. Whereas for a water quantity model the origin of the runoff is not important, such features show that water quality parameters must be traced more closely through a Lagrangian description. Since this cannot be done, we must expect a larger uncertainty for a WQM than we need to accept for a hydrologic water quantity model. At this time, we do not know how to handle such problems. The best we can do is to use models in conjunction with field measurements, in order to find the optimum model simplifications which yield meaningful answers to operational problems at acceptable costs.

Distributed Water Quality Models

"Distributed water quality models" are extensions of the physically-based strip models of figure 8 that are obtained by integration over larger and inhomogeneous areas. An example of such a model for small areas is the European SHE model (Beven et al. 1982, Beven 1985). Usually, they are based on partial differential equations of physics. They require for their use and calibration a very large amount of local information, which has to be obtained either by measurements, or by transferring data from other sites. Their advantage is that they permit the calculation of area distribution of quantities of the cycles of water and of pollutants. Whenever spatially variable input-output relations must be determined, such models are required. For water quality decisions,

they permit the calculation of pollutant paths, as influenced, for example, by changes in land use or by localized pollutant sources.

Natural variability for distributed WQMs. For distributed WQMs the model parameters are best inferred from extensive field experiments. Unfortunately, even more than for homogeneous area models the natural properties of a basin which influence the transport of pollutants are described by parameters which have large natural spatial and temporal variabilities, and what has been said above in the context of homogeneous area models holds for models for distributed WQMs also.

Models using distributed rainfalls and runoffs as inputs have been constructed; in some of these models the variability of the processes is considered. For example, one may estimate the exceedance probability of critical concentrations, such as in the model STREAM (Donigian et al. 1984) which estimates the exceedance probabilities for pesticide concentrations. It not only uses overland flow information, but it also incorporates hydraulic and chemical interactions.

Field programs for experimental support for this type of model are extremely costly and time consuming. Consequently, not many detailed field studies have been conducted. For example, 2in the Federal Republic of Germany (FRG), only few areas exist in which extensive measurements have been taken; the most important one lies in the fairly level loess area northwest of the Harz mountains near Brunswick (Walther 1980a and b). Therefore, parameters for physically-based models must usually be obtained from laboratory experiments, which may not apply to field situations: a well known example is the fact that the hydraulic conductivity of soils is highly dependent on the structure and quantity of macropores present. Sampling generally tends to destroy the macropore structure, or leads to samples which are too small to be representative of field soils.

Water Quality Models of Hydrologic Fields

In the most general type of model all processes are described by mathematical fields, usually employing partial differential equations. By means of a "field model" the processes of the water cycle and the resulting transport of pollutants is predicted at every point in time and space. The typical example for a (simplified) field model is the flow model for groundwater, in which the flow velocity and the pressure at every point in space is predicted. In theory, such a model could be free of randomness - it permits, in principle, to calculate for a given input field (of rainfall and pollutant, as well as ground cover and input of solar radiation, etc.) the concentration field in time and space, by means of partial differential equations which describe the physical, chemical and biological processes. In actuality, even the most elaborate models cannot be set up without recourse to simplification: nobody could determine all the necessary information which has to go into such a model. Because of such complexities, field models for a basin usually are simplified by separating surface flow from groundwater flow. In more elaborate models, these two horizontal flows are connected through an infiltration model with vertical flow, which appears as infiltration loss in the equations for overland flow, and with change in sign as recharge for the groundwater flow. Simplifications are obtained by eliminating one spatial dimension through integration: the groundwater equations are integrated over the depth, and the open channel flow equations over the cross-sectional areas. The simplification thus introduced leads to a deviation of the calculated results from the measured results, which usually is treated in the same manner as a measurement error and which is reduced by model calibration.

Process models of field type can profit from the possibilities of computer graphics (see Loucks et al. 1985, for recent references). Geographical Information Systems (GIS) have been developed which permit the user to create parameter maps with relative ease, and decisions and their consequences can be traced in their areal extent with great convenience. There is no question that these methods open a very exciting way of presenting calculation results, but it must be realized that they have their drawbacks: they are expensive to produce. The more realistic the graphical

output of a computer plotter looks, the more time it took to write the corresponding program. Also, the plots cannot overcome the basic shortages inherent in the models which produced them. Computer graphics can only present either measured data or model outputs, and both graphs can only be as good as their basis: the appearance of the graphical display gives the illusion that the calculated flows and their effects are an exact duplication of nature. Unfortunately, they sometimes appear to be more realistic than reality, although the incorporation of error bands obtained from a sensitivity analysis can help correct such an impression.

Another disadvantage of complex field models is that the number of cases which can be calculated with them is limited. This is an important shortcoming for WQMs which involve many process submodels, and often very complex flow models. It is therefore very likely that future developments will be in the direction of simplifying WQMs through the use of conceptual models, as has been done in surface hydrology.

Conceptual Water Quality Models

Historically speaking, hydrologists have tended to work on two levels. One level was to obtain an understanding of the physical processes involved in the hydrologic cycle, by looking at the processes at the scale of an area element. But they found out that area element information could not be transferred to whole basins. They bridged the gap between area elements and basins by introducing a second modeling level, that of "conceptual models", which is a term used to describe models having a rainfall function as input, and a runoff hydrograph as output. An extension to multidimensional outputs of water quality and quantity variables leads to the concept of a conceptual WQM.

The classical case of a conceptual model is the unit hydrograph model for describing a linear rainfall-runoff process, in which the rainfall is transformed into a runoff hydrograph by means of a one-dimensional convolution. The concept can be extended to composite models, in which many subbasins are linked to form the model for a large basin. Such models are used extensively in the FRG to determine the effectiveness of systems of flood-retention reservoirs (see Plate et al. 1988, for a description).

Water quality models of homogeneous basins. A homogeneous basin is an area with a fairly homogeneous cover of fields, dwellings, and forests, and a topography which is not too uneven. In the hydrologic models used in the FRG, the sizes of homogeneous basins range from a few to about 10 km^2. For such basins, the unit hydrograph theory has been developed. The unit hydrograph is a function representing the catchment response by means of which the (point-) rainfall is converted into a runoff hydrograph through a convolution

$$Q(t) = \alpha A \int_0^{\infty} h(\tau)\, i(t-\tau)\, d\tau \qquad [10]$$

where α = a conversion factor for the units used,
$A\ [L^2]$ = the catchment area,
$h(t)\ [T^{-1}]$ = the unit hydrograph as function of time, and
$i(t)\ [L/T]$ = the effective intensity of rainfall averaged over the basin area.

When the unit hydrograph cannot be measured directly because of lack of data, a model for regionalization is often used, such as the model developed by Lutz (1984). This model consists of a part for predicting the runoff coefficient, and a second part for the prediction of the unit hydrograph's peak value and for its concentration time (time to peak). Both quantities have been correlated with total rainfall, basin area, percentage of area covered by forests, and percentage of area covered by houses, for more than 70 small basins, and for more than 600 measured rainfall-runoff hydrographs. Needless to say that all these quantities are subject to substantial

uncertainties, and it is one of the foremost tasks faced by hydrologists to quantify these uncertainties and to assess what effect they have on engineering decisions.

In analogy to the unit hydrograph, pollutants may be routed through an impulse response function, or "pollutograph", $h_s(t)$ in analogy to the unit hydrograph h(t) of runoff (Rinaldo and Marani 1987, Jury et al. 1986). The simplest shape of $h_s(t)$ is obtained if we assume proportionality between rainfall and pollutant, so that $Q_s(t) = aQ(t)$. However, other cases of "pollutograph" are also useful: it is often noticed, for example, that due to the processes described above the peak of the pollutant wave arrives faster than that of the flood wave, which would imply a unit pollutograph with a shorter rise time than the flood unit hydrograph.

In general, unit hydrographs and "pollutographs" are used deterministically, i.e. a deterministic input function is used in conjunction with a deterministic h(t) function. However, it is not difficult to use a deterministic h(t) with random event inputs, by employing simulation techniques. A special case is a random input consisting of a time series with normally distributed input magnitudes; in this case the output also is a normally distributed variable, with a variance which is determined through the spectrum function of the input process and the Fourier transform (or system function) of h(t). Some use of these relations have been made for groundwater models; see for example Geldner (1981).

Simulation also permits to incorporate uncertainties, and the result of the calculations is the pdf of the water quality outflow of the system considered.

Water quality models of composite hydrologic basins. Recent practice in the FRG (and elsewhere) has been to subdivide large hydrologic basins into subbasins. For each of these basins, a unit hydrograph is developed, either from direct measurements or by means of a regional model. The runoff hydrographs from the subbasins are linked by a network of channels, through which the hydrograph is routed by using techniques which range from pure translation to nonstationary calculations by means of the complete St. Venant equations.

An extension of the model for sediment yield has been the combination of the USLE or MUSLE with conceptual, or basin-type models. Although there does not yet exist a model in which pollutants have been included into a hydrologic model of such a type, such models were used for calculating the rate of erosion from soil surfaces and deposited in reservoirs (Bogardi et al. 1985, Hrissanthou 1986). Both authors used simulation methods with historical rainfall inputs to obtain the historical sequence of event-based erosion yields, which were then summed over the year. For the model the basin was subdivided into many small, approximately field-size subbasins, to each of which the modified Universal Soil Loss Equation (MUSLE) has been applied. An annual precipitation index K was used, but event dependence was introduced by the use of a daily value of precipitation from which the sediment inflow to the reservoir was calculated for the daily effective rainfalls obtained from historical records. The differences between the two studies were in the way the MUSLE was used for the subbasins, and how they were linked for the total yield. For each partial area the assumptions made by Hrissanthou (1986) were about the same as for the model of Smith et al. (1977), which was described above. However, instead of using the SCS method for determination of the effective rainfall, he used a modification of the SCS method developed by Lutz (1982). Furthermore, Hrissanthou included a channel sediment routing subroutine (Williams 1975). Both Hrissanthou and Bogardi et al. compared their results with measurements of annual reservoir deposition. Good agreement was found for average annual values, but considerable scatter existed in daily sediment yields.

The incorporation of stochasticity in the form of natural variability and uncertainty proceeds as has been described for models of homogeneous basins. However, no reports are available in which such an investigation has been performed and compared with field results. Since this approach

appears to be most promising for large-scale pollution problems, it is likely that this situation will change in the near future.

STOCHASTICITY AND DECISION MODELS OF WQMs

Because the general problem illustrated in figure 2 is extremely complex, simplifications are in order if design or management decisions must be made. Decision models have components with outputs that are modifiable by means of decision variables, i.e. those variables which can be influenced directly or indirectly by human actions. The type of model to be used in water quality decision processes must depend on the decisions which are contemplated. For example, a model designed to predict the peak concentrations of pesticides in a river at a particular point can have a different structure than a model which predicts the amount of pesticide which penetrates into the groundwater in a subarea of the basin. Furthermore, because uncertainty of the different kinds considered may limit the quality of the information obtainable from a model, it is useful to consider the tradeoff of model uncertainty against natural variability. Obviously, if the variance of the pdf for the model uncertainty is small compared to that of the sample or the measurement, then it is not worthwhile to further improve the model.

More generally, the value of a WQM can best be assessed if a numerical or at least an ordinal preference value is assigned to the consequences of a decision which is based on the model. If a clearly advantageous decision can be made which is independent of the quality of the model, then there is no operational sense in further improving the model. If, on the other hand, costly investments would be necessary for preparing against possible but uncertain consequences, then model improvements, or improvements in the data basis may be in order. The generalized risk, as defined in Duckstein et al. (1987) which quantifies such decisions, is therefore an important figure of merit for the value of WQMs.

Risk as Figure of Merit

In the context of stochasticity of decision models we define, in the sense of decision theory (Berger 1985) risk as follows. Let $\mathbf{y} = (y_1, y_2, ..., y_I)$ be those variables from the output vector of a WQM which can be manipulated by decisions, so that their values are conditional on the decisions $\mathbf{d} = (d_1, d_2, ..., d_j)$ ($\mathbf{d}$ = vector of decision variables). Let these variables occur in combinations determined by the joint probability density function (joint pdf) given by:

$$f(\mathbf{y}|\mathbf{d}) = f(y_1, y_2, ...y_I, ...d_1, d_2, ..., d_j). \qquad [11]$$

Furthermore, let $K(\mathbf{y}|\mathbf{D})$ be the function which describes the consequences of the occurrence of the combination y_1, y_2 ... for a given decision vector $\mathbf{d} = \mathbf{D}$. Then the risk is defined in general as the expected value of K over the (conditional) pdf $f(\mathbf{y}|\mathbf{D})$, or:

$$RI(\mathbf{D}) = \int_{-\infty}^{\infty} K(\mathbf{y}|\mathbf{D})\, f(\mathbf{y}|\mathbf{D})\, dy \qquad [12]$$

where the integration has to be performed over all the elements of the vector $\mathbf{y}$.

The risk is a measure of the gamble which we are taking, if we make the decisions D. It is a single-valued number, - called a Figure of Merit, FM (Duckstein et al. 1987), which permits to judge the value of the decision. Note that there might be other figures of merit associated with any one decision process. FMs may be based on different criteria than those which can be calculated from the WQM, or which are based on a limited number of output variables, rather

than on the complete set. Also, different types of risk may occur, depending on the definition of the consequence function. For details reference is made to Duckstein et al. (1987).

To illustrate this concept, consider the case of a single output variable y with pdf $f(y|\mathbf{d}) = f(y)$ (i.e. we assume that y is stochastically independent of **d**), with a consequence function $K(y|\mathbf{d}) = K(y|\mathbf{d})$ (i.e. **d** is assumed to be a single variable). As an example, assume that $y_1 = z$ be the inflow of eroded soil into a lake, with probability density function f(z). With the eroded soil, insecticide is carried in a quantity pz where p is the decision variable proportional to the amount of pollutant applied to the soil. The consequence function $K(z|\mathbf{p}) = K(zp)$ denotes the damage done to the ecology of the lake. Then the risk may be taken as the expected value of the damage done to the lake: In a K(.) vs. z plane, $K(z|\mathbf{p})$ is a different curve for each **p**, as is indicated schematically in figure 10. The expectation:

$$RI(\mathbf{p}) = E\{K(z|\mathbf{p})\} = \int_{-\infty}^{\infty} K(z|\mathbf{p})\, f(z)\, dz \qquad [13]$$

is a function of **p**. Note that this expression only reflects the natural variability. It does not account for uncertainty - this will be illustrated next.

Note also that equation 13 only quantifies the consequences of the decision through the risk, which is a decision model input, not its result. The extension of this model to a decision situation, where an optimum decision has to be formulated, is beyond the scope of this paper (see Berger 1985).

Uncertainty in Decision Models

In general, a WQM is also subject to uncertainty. Uncertainty has the effect that the true risk in our gamble against nature is not known. Uncertainty causes RI to become a random variable,

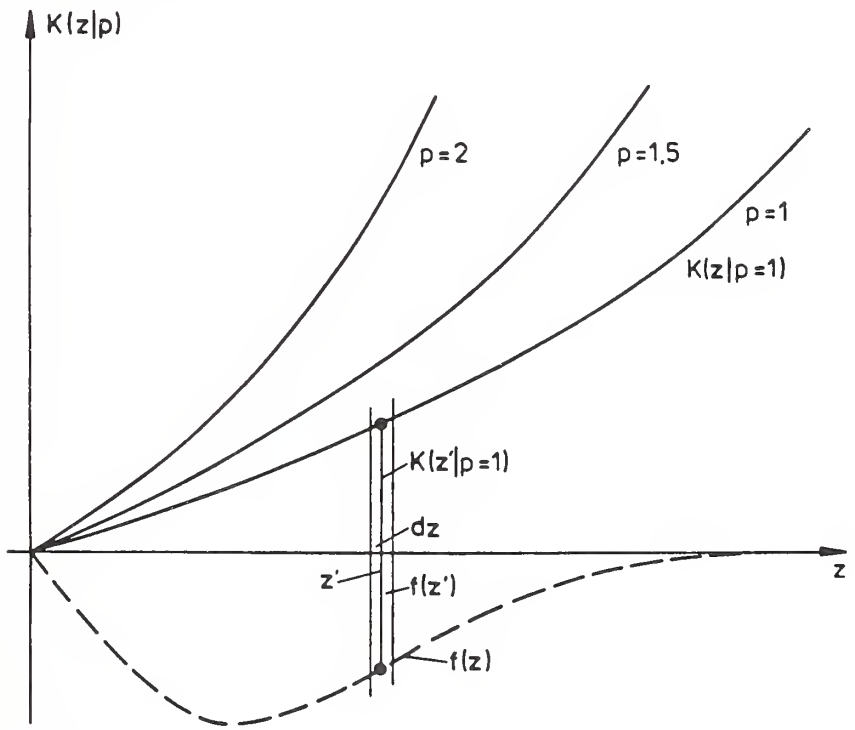


Figure 10.
Display of integration for risk RI(p).

which we shall call goal function g(.), instead of a single number for each **D**. The goal function is a random variable having a conditional pdf f(g(.)|**PAR**) which must be estimated from the uncertainty of the parameter vectors, **PAR**, which affects the WQM. The best estimator for RI(D) is the expectation of g(.).

As an illustration, consider the case of figure 10. We see that uncertainties in models or parameters will cause $f(t_r, P_r)$ to change its shape, and thus f(z) also becomes a variable quantity. The variability can be expressed by replacing f(z) by the conditional pdf f(z|**PAR**). Assume that the variability of the parameter vector **PAR** can be accounted for through the probability density function, f(**PAR**). Then the goal function becomes:

$$g(d,\mathbf{PAR}) = \int_{-\infty}^{\infty} K(d,z|\mathbf{PAR})\, f(z|\mathbf{PAR})\, dz. \qquad [14]$$

Note that the goal function has a pdf depending on the pdf f(**PAR**) of the parameters. From g(d,**PAR**) the best estimate for the risk RI is obtained as expectation of g(.) over f(**PAR**), which is called the Bayes risk, BR(d):

$$BR(d) = \int_{-\infty}^{\infty} g(d,\mathbf{PAR})\, f(\mathbf{PAR})\, d\mathbf{PAR}. \qquad [15]$$

which is the quantity to be used in decision models, if parameter uncertainty has to be accounted for. Again, for further developments reference is made to the literature.

CONCLUSIONS

In this paper we have given a brief survey of uncertainty and risk in water quality modeling. The subject is not simple, and it has not been possible to give more than an introduction. In fact, even the applications of well known statistical concepts to water quality models has not progressed very much, mainly because there exist only very few problems for which a complete closed-form analysis is possible, and most solutions will have to be obtained by simulation. However, we feel that a statistical approach (or perhaps a fuzzy analysis) is the only sensible approach to many important water quality problems, because it will hardly ever be possible to find enough data, and to perform long enough calculations, for obtaining all the information which is required for decision-making in increasingly complex environmental situations. The adoption of the statistical approach requires us to learn to think in different terms than in the past: for WQMs we must learn to be content with probabilistic answers instead of exact numbers. For this we must learn, in situations involving natural variability and uncertainty, to balance model or parameter accuracy against this natural variability.

REFERENCES

Beck, M.B. 1987. Water quality modeling: A review of the analysis of uncertainty. Water Resour. Res. 23:1393-1442.

Berger, J.O. 1985. Statistical decision theory and Bayesian analysis. Springer, New York.

Beven, K.J., R. Warren, and J. Zaoui. 1982. SHE: Towards a methodology for physically based distributed forecasting in hydrology. In Hydrological Forecasting, IAHS Publication No. 129, pp. 133-137.

Beven, K.J. 1985. Distributed models ch. 13 In Hydrological Forecasting, M.G. Anderson and T.P. Burt, eds., pp. 405-435, J. Wiley and Sons. NY.

Bogardi, I., and L. Duckstein. 1978. Input for a stochastic control model of P-loading. Ecol. Model. 4:173-195.

Bogardi, I., A. Bardossy, and L. Duckstein. 1985. Effect of parameter uncertainty on calculated sediment yield. Adv. Water Resour. 8:96-101.

Boulier, J.F., and M. Vauclin. 1984. Stochastic simulation of water flows in unsaturated soils. Comparison with observed values. In P. Udluft et al. eds., Proceedings, International Symposium on recent investigations in the zone of aeration, pp. 315-330, Munich.

Bresler, E., and G. Dagan. 1983. see Dagan and Bresler (1983) (below).

Chow, V.T., ed. 1964. Handbook of Applied Hydrology McGraw-Hill, New York.

Cordova, J.R., and I. Rodriquez-Iturbe. 1985. On the probabilistic structure of storm surface runoff. Water Resour. Res. 21:755-763.

Cordova, J.R., and R.L. Bras. 1981. Physically based probabilistic models of infiltration, soil moisture and actual evapotranspiration. Water Resour. Res. 17:93-106.

Dagan, G., and E. Bresler. 1983. Unsaturated flow in spatially variable fields. 3 papers, Water Resour. Res. 19:413-435.

Donigian, A.S. Jr., D.W. Meier, and P.P. Jowise. 1984. Stream Transport and agricultural Runoff of Pesticides for Exposure Assessment: A Methodology Environmental Research Laboratory US Environmental Protection Agency. Athens, Georgia.

Duckstein, L., M.M. Fogel, and C.C. Kisiel. 1972. A stochastic model of runoff producing rainfall for summer type storms. Water Resour. Res. 8:410-421.

Duckstein, L., I. Bogardi, and M. Fogel. 1978. An event based stochastic model of phosphorus loading into a lake. Adv. Water Resour. 1:321-329.

Duckstein, L., E.J. Plate, and M. Benedini. 1987. Water engineering reliability and risk: a system framework. In L. Duckstein, and E.J. Plate eds., Engineering Reliability and Risk in Water Resources. NATO Advance Study Series, Martinus Nijhoff. pp. 1-20.

Geldner, P. 1981. Deterministische und stochastische Methoden zur Bestimmung der Selbstdichtung von Gewässern Mitteilungen No.49, Institute for Hydraulic Engineering, Stuttgart University, FRG.

Hann, C.T. 1977. Statistical methods in hydrology. Iowa State University Press, Ames.

Haith, D.A., and L.J. Tubbs. 1981. Watershed loading functions for nonpoint sources. Proc. ASCE, J. Environ. Eng. Div. 107:127-137.

Hrissanthou, V. 1986. Computation of sediment yield from a large watershed in middle Europe. Proceeding, Third Internatl. Symposium on River Sedimentation, Jackson, Miss.

Jacks, G., G. Knutsson, L. Maxe, and A. Fylkner. 1984. Effect of acid rain on soil and groundwater in Sweden. Chapter 7 In B. Yaron et al. ed., Pollutants in Porous Media. pp. 94-114, Springer, New York.

Julien, P.Y., and A.M.Dawod. 1987. On predicting upland erosion losses from rainfall depth. Stoch. Hydrol. Hydraul. 1:127-134.

Julien, P.Y., and M. Frenette. 1985. Modeling of rainfall erosion. Proc. ASCE J. Hydraul. Eng. 111:1344-1359.

Kelman, S., W. Storch et al. 1977. ASCE Task Committee on Agricultural Runoff and Drainage, Report on Quality aspects of agricultural runoff and drainage. Proc. ASCE, J. Irrig. Drain. Div. 103:475-495.

Loucks, D.P., J. Kindler, and K. Fedra. 1985. Interactive water resources modeling: An overview. Water Resour. Res. 21:95-102.

Lutz, W. 1984. Berechnung von Hochwasserabflüssen unter Anwendung von Gebietskenngrössen Reports No. 24, Institute for Hydrology and Water Resources Planning, University of Karlsruhe, FRG. (see also: Plate et al. 1988)

Maniak, U. 1986. Rainfall runoff model for rural catchments. Proceedings, Fourth International Hydrology Symposium: Multivariate Analysis of Hydrological Data. Water Resources Publications, Littleton,Colo. pp. 880-893.

Milon, J.W. 1987. Optimizing non-point source controls in water quality regulation. Water Resour. Bull. 23:387-396.

Plate, E.J. 1976. Intermittent processes and conditional sampling. In Gottschalk et al. eds., Proc. Second IAHR Symposium on Stochastic Hydraulics, Water Resources Publications, Colorado, pp. 3-29.

Plate, E.J. 1986. Multivariate analysis of rainfall models. Proceedings, Fourth International Hydrology Symposium. Multivariate Analysis of Hydrological Data. Water Resources Publications, Littleton, Colo. pp. 66-87.

Plate, E.J., J. Ihringer, and W. Lutz. 1988. Operational models for flood calculations. Proc. IAHR Congress of the Asian Pacific Division, Kyoto.

Rosso, R. 1986. Time series analysis and stochastic models of sedimentation processes at different scales. Proceeding, Third Internatl. Symposium on River Sedimentation, Jackson, Miss. pp. 1707-1720.

Smith, J.H., D.R. Davis, and M. Fogel. 1977. Determination of sediment yield by transferring rainfall data. Water Resour. Bull. 13:529-541.

Smith, R.E., and R.H.B. Hebbert. 1983. Mathematical simulation of interdependent surface and subsurface hydrological processes. Water Resour. Res. 19:987-1001.

Sorooshian, S., and V.K. Gupta. 1983. Automatic calibration of rainfall runoff models: the question of parameter observability and uniqueness. Water Resour. Res. 19:260-268.

Szidarovsky, F.S., F.S. Yakowitz, and R. Krzysztofowicz. 1976. A Bayes approach for simulating sediment yield. J. Hydrol. Sci. 3:33-45.

Todorovic, P., and E. Zelenhasic. 1970. A stochastic model for flow analysis. Water Resour. Res. 6:1641-1648.

Troutman, B.M. 1985. Errors and parameter estimation in precipitation - runoff modeling: 1. Theory. Water Resour. Res. 21:1195-1213.

Walling, D.E. 1977. Assessing the accuracy of suspended sediment rating curves for a small basin. Water Resour. Res. 13:531-538.

Walther,W. 1980a. Prozess des Stoffabtrags und der Stoffauswaschung während und nach Starkregen in ackerbaulich genutzten Gebieten: 1.Bericht: Stoffabtrag. Zeitschrift f. Kulturtechnik und Flurbereinigung 21:65-74.

Walther, W. 1980b. Prozess des Stoffabtrags und der Stoffauswaschung während und nach Starkregen in ackerbaulich genutzten Gebieten: 2.Bericht: Stoffauswaschung. Zeitschrift f. Kulturtechnik und Flurbereinigung 21:145-153.

Waymire, E., and V.K. Gupta. 1981. The mathematical structure of rainfall representation: 1.A review of the stochastic rainfall models. Water Resour. Res. 17:1261-1272.

Williams, J.R. 1975. Sediment yield predictions with universal equation using runoff energy factor. Agricultural Research Service. ARS-S-40, USDA, Wash. pp. 244-252.

Woolhiser, D.A., and H.B. Osborn. 1985. A stochastic model of dimensionless thunderstorm rainfall. Water Resour. Res. 21:511-522.

Yevjevich, V. 1984. Structure of daily hydrological series. Water Resources Publications, Littleton, Colo

DISCUSSION OF THE PAPERS PRESENTED IN TECHNICAL SESSION 6, PART 3: RISK ANALYSIS/CONFIDENCE LIMITS

R.J. Hanks[1], Presiding
L.S. Willardson[2], Recorder

PAPERS DISCUSSED

Analyzing Statistical Properties of Nonpoint-Source Water Quality Variables by J.D. Salas, J.C. Loftis

Stochastic Aspects of Water Quality Modeling for Nonpoint Sources by E.J. Plate and L. Duckstein

SPECIFIC QUESTIONS AND COMMENTS

Question: (D. Jackson, Susquehanna River Basin Commission, Harrisburg, Pennsylvania) Relating to a stochastic modeling problem treated in my poster paper, how can stochastic modeling be done for data that are not equally spaced in time?

Response: (J. Loftis, Department of Agricultural and Chemical Engineering, Colorado State University, Fort Collins, Colorado) The time scale could be enlarged so that each time step would have a representative data point. For example, use quarterly time steps instead of monthly time steps. The realities of data sets limit the use of time series approaches.

Question: (R. Hanks, Soils and Biometeorology, Utah State University, Logan, Utah) For the types of analyses suggested to design monitoring systems, is it not true that the data is required before the analysis can be made?

Response: (J. Loftis) The problem is similar to the chicken and the egg. However, known physical relationships or statistical rules-of-thumb can be used to design monitoring systems that will give the required data. Data collection programs should often start small. They can be enlarged, modified and refined as time progresses. Some suggestions for monitoring system design are given in the written paper.

Comment: (E. Plate, Institute Fur Hydrologie, Federal Republic of Germany) Time series are useful for processing data that will be used as input to models developed or based on known physical relations. Unit hydrographs can be combined with rainfall data and known transfer processes to design monitoring systems. The availability of computers makes possible the generation of models that can use known physical relationships. Water quality problems can be examined on the basis of physical principles and then historical data can be used for validation.

Comment: (J. Loftis) The kind of modeling represented by ARMA-type models that were in vogue ten years ago has not been generally accepted in water quality studies because of data limitations. It is often easier to develop data over a short time for more physically realistic models. For regulatory purposes, very simple statistical methods may be appropriate. Time-series modeling has application where long water quality data records are available and some suggestions are made in the paper on which methods might be appropriate.

[1]R.J. Hanks, Professor, Soils and Biometeorology, Utah State University, Logan, Utah.
[2]L.S. Willardson, Professor, Department of Agricultural/Irrigation Engineering, Utah State University, Logan, Utah.

Question: (D. Gustafson, Monsanto, St. Louis, Missouri) Why has there been a recent increase in the nitrate problem mentioned in papers from Nebraska and Iowa? Do you think reported increases are real or is it possible that they are just doing more measuring?

Response: (J. Loftis) The additional measurements may have some effect, but there are some radioisotope studies that indicate that the agricultural application has caused increases in the nitrate content of water supplies. With pesticides, of course, we are sure that reported occurrences are due to human activities.

Comment: (W. Kinzelbach, Institute Fur Wasserbau, University of Suttgart, Federal Republic of Germany) Uncertainty in models is not important if the uncertainty is stable. A problem arises when small changes in inputs cause large changes in outputs. This type of model behavior is very dangerous where models are used to extrapolate to conditions radically different from those used in development.

Response: (E. Plate) The process function is critical. It is always risky to extrapolate them, but if they are exactly known, their effect shows up dependably in the output. It is sometimes possible, in complicated models, to trade accuracy for uncertainty and obtain useful results. A general model can be used with a range of acceptable random variables to check the sensitivity of the model. Then a statistical analysis can be used to tell whether the model results can be extrapolated.

Question: (H. Roaza, Florida Department of Environmental Regulation, Tallahassee, Florida) Even though more data is assumed to be better, is it not possible to collect biased data? For example, collecting all groundwater data in a high-risk area and leaving the low-risk area without data. Limited resources results in this problem.

Response: (E. Plate) If the consequence function in a model is wrong, the statistics may be good and the results will still be unacceptable. Models should be examined in light of what is required. If high-risk-area models are of interest, then high-risk area samples are needed and are appropriate. The input model, the process model, and the output or consequence model need to be examined for their relative importance with respect to the application problem. Do not use a complicated model if there is only a need to know whether some factor is high or low.

Comment: (J. Loftis) Modelers should define what they are looking for. Sometimes a stratified random sampling approach is needed initially to assure that statistical conclusions describe the population which is really of interest.

Question: (G. Oliver, Dow Chemical Company, Midland, Michigan) In designing sampling programs, how is the decision made about how much variability or results to allow?

Question: (S. Glasser, USDA Forest Service, Atlanta, Georgia) How much error can be allowed?

Response: (J. Loftis) Decision-makers or managers generally do not know what confidence intervals are appropriate in a new study. However, a model might be used to evaluate the confidence intervals provided by different levels of monitoring. Managers can then choose between various monitoring alternatives based on their anticipated level of performance.

Response: (E. Plate) In Germany, according to new law, potential polluters are asked to set the amount by which they will exceed legal water quality standards. For exceeding, they pay a fee or fine. The regulatory agencies then just take samples to determine if the self-imposed pollution levels have been exceeded. The company has to meet its own standard or it is fined rather heavily.

Comment: (J. Troiano, California Department of Food and Agriculture, Sacramento, California) Data sets are demanded by administrators, but it may be dangerous to use data sets we have in hand for purposes other than those for which the data are collected. For example, sampling soils of the same surface type showed large differences in subsurface characteristics that could seriously affect model results on pollutant movement.

Response: (J. Loftis) A connection is needed between the decisions and the type of data needed for those decisions. Many decisions are currently made based on data that do not support them. However, we are gradually moving in the direction of defining the types of data needed to make particular decisions.

ROOT ZONE PROCESSES AND WATER QUALITY: THE IMPACT OF MANAGEMENT.

B.E. Clothier[1]

ABSTRACT

Better management to minimize the burgeoning problem of agricultural non point-source pollution will arise from clearer understanding of the small-scale processes of water and chemical movement in the root zone of crops. As a precursor, improved field measurement and analytical description will be required. Four critical processes are identified as being poorly understood.

(1) The effect of aerial plant parts on controlling the arrival of water and chemical at the soil surface.
(2) The role of the structure of the soil surface in prejudicing the pattern of water and chemical entry into the root zone.
(3) The dynamics and behavior of plant roots as they extract water and chemical.
(4) The unsaturated redistribution of water within and below the root zone.

Computer simulation is one means by which these small-scale processes can be synthesized and integrated into models for assessing the impact of agricultural management on water quality. However it is imperative that these mechanisms are correctly described, appropriately parameterized and rationally incorporated in any simulation model intended for use in water quality assessment on a regional scale.

INTRODUCTION

The roots of agricultural, horticultural and forest plants ramify some fraction of the fertile surface soil in order to exact support and extract sustenance in the form of water and nutrients. Conversely plants substantially modify the surface soil not only by virtue of their physical presence, but also because of the myriad edaphic processes they engender. Mankind, by cultural management of plants, attempts to optimize the harvestable yield of either their skeletal, vegetative or floral parts. For example, management by irrigation is commonly used to provide that amount of water not easily forthcoming from the soil water reservoir of the root zone. Frequently, fertilizers are used to enhance the growth of plants to a level that would not naturally be possible because of limited nutrient supply in the root zone. Pesticides are often applied to crops to reduce the depredations of pathogens, or to eliminate competition from other species.

Effective management of any applied water, and many agricultural chemicals, requires that they become rapidly distributed in sufficient concentration to be easily available over a requisite portion of the root zone. Unfortunately the lability of these applied chemicals, while essential for productive plant growth, renders any unused fraction liable to export from the root zone, often with deleterious environmental consequences.

The form of land use, and its management, can substantially alter the balance between the inputs and losses of applied agricultural chemical. Both naturally-fixed and urine nitrogen are essential for the productive growth of pasture in New Zealand and elsewhere. Fertilizer nitrogen is often required to ensure profitable production of horticultural tree and vegetable crops, as shown in

[1]B.E. Clothier, Scientist, Plant Physiology Division, D.S.I.R., Palmerston North, New Zealand.

table 1. Patently the nature of land use and management can affect the quality of water emanating from tracts of farmland.

The four-quadrant modification by de Willigen and van Noordwijk (1987) of de Wit's (1953) scheme for analyzing the nutrient response of crops will be used here to provide a conceptual basis for assessing the effect of management on water quality. This scheme illustrated in figure 1, can be broadened to apply beyond fertilizer. The chemical could simply be water, or even a plant-systemic compound used in disease or pest control. Even agricultural drainage can be viewed from this perspective; the abscissa of quadrant II would become chemical removed, namely drainage water. The ordinate of quadrants III and IV would then be the availability of the complement of soil water, namely oxygen. The intimate connections between management, the efficacy of agricultural chemical use, and the various scientific disciplines are economically highlighted by this structure.

An agronomic goal has been to establish productive management strategies solely by determining the response of crop yield to applied chemical (quadrant II), whether the latter be fertilizer, simply water, or indeed pesticide. With understandable myopia this has focused on optimization of harvestable yield, guided by the law of diminishing returns. Little cognisance can be given to process. These data are obtained from replicated field trials and inferences obtained simply by analysis-of-variance procedures. Mass balance considerations are only indirectly referenced by noting that the efficiency of the chemical lessens as the application rate increases. Meanwhile at a much reduced length-scale, and often under ad libitum conditions, plant physiologists study the consumption of water and the incorporation of nutrients into plant organs (quadrant I). By way of contrast, quadrants III and IV tacitly acknowledge the balance of mass. The possibility of loss of chemical within and beyond the root zone is recognized. There may be an initial pool of available chemical in the root zone which will produce the basal yield from the minimum plant uptake. Management may seek to lift crop yield via intensification and increased application of chemical. Additional yield is brought about by a greater amount of chemical being made available in the root zone. However some increasing fraction might not be available as it may either be volatilized, evaporated, become fixed by the soil, change its chemical state, or be lost by leaching (quadrant III). Soil scientists study these processes in order to assess the magnitude of plant uptake vis a vis loss through 'unavailability'. While 'availability' is a necessary condition for plant growth it is not sufficient (quadrant IV). The roots must absorb the chemical for it to be of systemic benefit. Failure of roots to ramify the soil completely will result in available chemical being left and liable to loss in the environment. Definition of 'completely' is dependent upon the interaction between chemical mobility and the growth and density of the root absorption surface. Root ecologists are struggling with the task of describing the topological complexity of root density and branching (Gandar and Hughes 1988, de Willigen and van Noordwijk 1987, Mandelbrot 1982) in relation to the pattern of chemical uptake (Böhm 1975, Nye and Tinker 1977).

Quadrants III and IV illustrate that as application increases, efficiency declines. Reciprocally, environmental anxiety rises! The leaching loss of one chemical, nitrogen, is shown in table 1 for various forms of land use in New Zealand. The losses are exacerbated by the coincident application of another agricultural chemical- irrigation water. Unfortunately the level of nitrate-N commonly deemed to be acceptable in drinking water (11.3 mg N03-N/l) appears incompatible with many forms of productive agriculture (O.E.C.D. 1986, Hubbard et al. 1986).

Rijtema (1987) fears that even a "... reduction of fertilization to a sub-optimal level ... does not offer a solution in all cases". Nevertheless de Willigen and van Noordwijk (1987) suggest that achievement of any increase in efficiency, to realize a reduction of loss to the environment, has to be based on improved efficiencies in quadrants III and IV. Modeling the physical, biological and chemical processes operating in the root zone is a research imperative. The modeling task will be challenging and scientifically daunting. But worse the public have an unrealistic estimate of how

Table 1.
An assessment of the nitrogen inputs and leaching losses from different forms of land use in New Zealand (from Rutherford et al. 1987, Burden 1982, Gandar and Ball 1982).

Input	Rate (kg-N/ha/yr)
Rainfall	2 - 5
Ryegrass/clover nitrogen fixation	35 - 350
Urine from extensive sheep grazing	40 - 70
Urine from irrigated sheep grazing	100 - 150
Horticultural fertilizer	120 - 150

Leaching Losses	Rate (kg-N/ha/yr)
Extensive sheep grazing	5 - 25
Irrigated sheep grazing	70 - 100
Plowed cropland	60 - 90
Irrigated horticulture	100 - 200

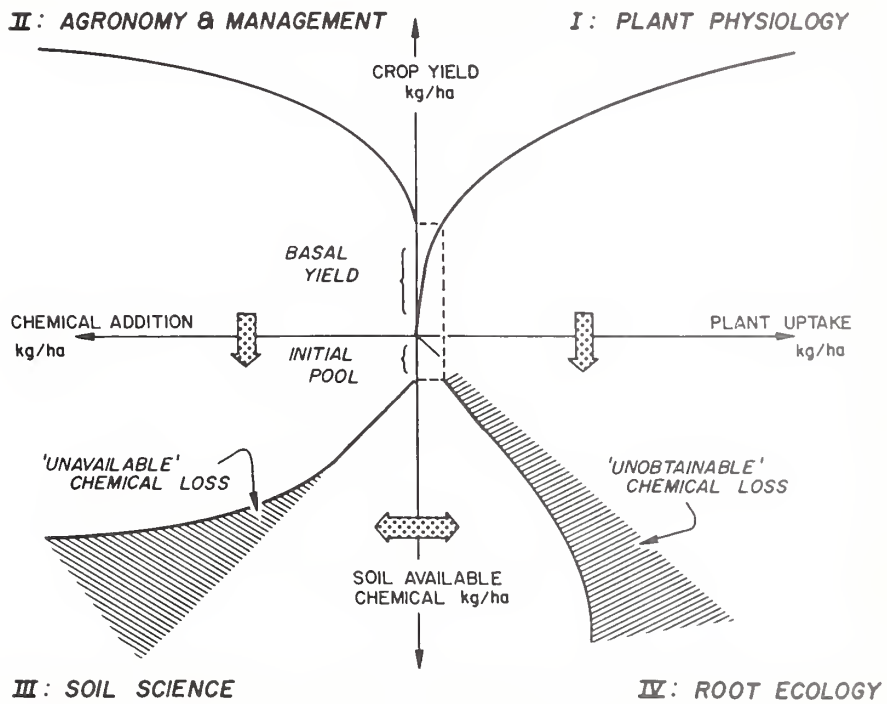


Figure 1.
The four-quadrant scheme of de Willigen and van Noordwijk (1987) for analyzing the response of crops to the application (or removal) of agricultural chemicals. The shaded areas indicate potential losses of chemical to the environment.

quickly and effectively the problems can be addressed and ameliorated (Garfield 1986). This is compounded by decades of adherence to the belief that groundwater was immune to contamination, on the erroneous tenet that the soil would bind chemicals and cleanse the water as it percolated through the unsaturated zone (Sun 1986).

MODELING

Modeling is sensu stricto the basis of all science: from observation and measurement the scientist deduces some abstract formal view of a particular process. This abstract realization whether it be conceptual, diagrammatic, analytical or numerical, is in essence a model. Recently however the democratic use of computers has tended to cause the terms 'modeling' and 'computer simulation' to become synonymous. Reactionaries would say this overwhelming of modeling by computer simulation may not only ruin science (Post 1974), it might even represent a threat to mankind (Truesdell 1984)! Others underwhelmed, merely suggest that computer simulation leads to forms of intellectual dishonesty, disguising of simplistic ideas, slip-shod work, space-age jargon and spurious mathematization (Andreski 1972, Philip 1975).

Simulation by computer does offer one means by which analytical results can be synthesized. The first point to be made about issues of water quality and non point sources is figure 2 (left). It is around the single, point source (or sink) on a length scale of the order of about lm, that soil physicists (quadrant III), root ecologists (quadrant IV) and plant physiologists (quadrant I) have successfully concentrated their analytical skills. For example, Clothier and Scotter (1982) studied irrigation water flow from the point source of a drip emitter; Kirkham and Powers (1972) gave solutions for water flow to line and plane-sink drains; Gardner (1960) analyzed water flow through soil to a line-sink root, and Monteith (1965) described the process of transpiration of leaf water from a stomatal orifice. But all this preoccupation is with the point Philip (1975) laments, and it has left unanswered "... how we may best use our understanding of small-scale local processes on the larger(r) scale ...the beautiful economy of analytical scientific methods is soon lost in the sheer magnitude, complexity and imprecision of the task of synthesis". When addressing the issue of water quality associated with a non point source (figure 2, right), the task of synthesis and integration of the points (figure 2, left) is enormous. This is especially so when we seek to assess the impact of management of large tracts of land on the quality of adjacent waters.

Scientific analysis has already revealed the pattern of nitrogen volatilization from, and nitrate movement under individual urine spots on pasture. This has been realized by way of controlled field experiments with simulated urination. The local intensity of application may be as high as 1000 kg-N/ha per urination (Ball and Ryden 1984, cf. table 1). Apparently some 15 percent (Ball et al. 1979) to 50 percent (Field et al. 1985) of the applied urine-N might be lost by leaching under a single urine spot. However we are less clear as to the issue of nitrogen by leaching, or surface run off, from the non point source that comprises the multitude of urine point sources whose distribution is governed by the idiosyncratic grazing habits of sheep or cattle (figure 3). Undoubtedly computer simulation can help with the tasks of integration and synthesis, but we should retain a healthy skepticism.

As maximization of crop production has long been the goal of agricultural management, not surprisingly crop growth simulation has been at the forefront of agricultural computer modeling. Can these models be useful for assessing issues of non point water quality? Crop simulation models do provide a framework for integrating research results for they are invariably hierarchical in nature. Characteristically they possess a mainline crop-growth routine (quadrant II) into which the machinations of various subsystems are fed; root growth and chemical uptake (quadrant IV), soil character and behavior (quadrant III), light interception with $C0_2$ and H_20 exchange (quadrant I), etc. The gross output of the model, crop yield, can be tested. But frequently success is only

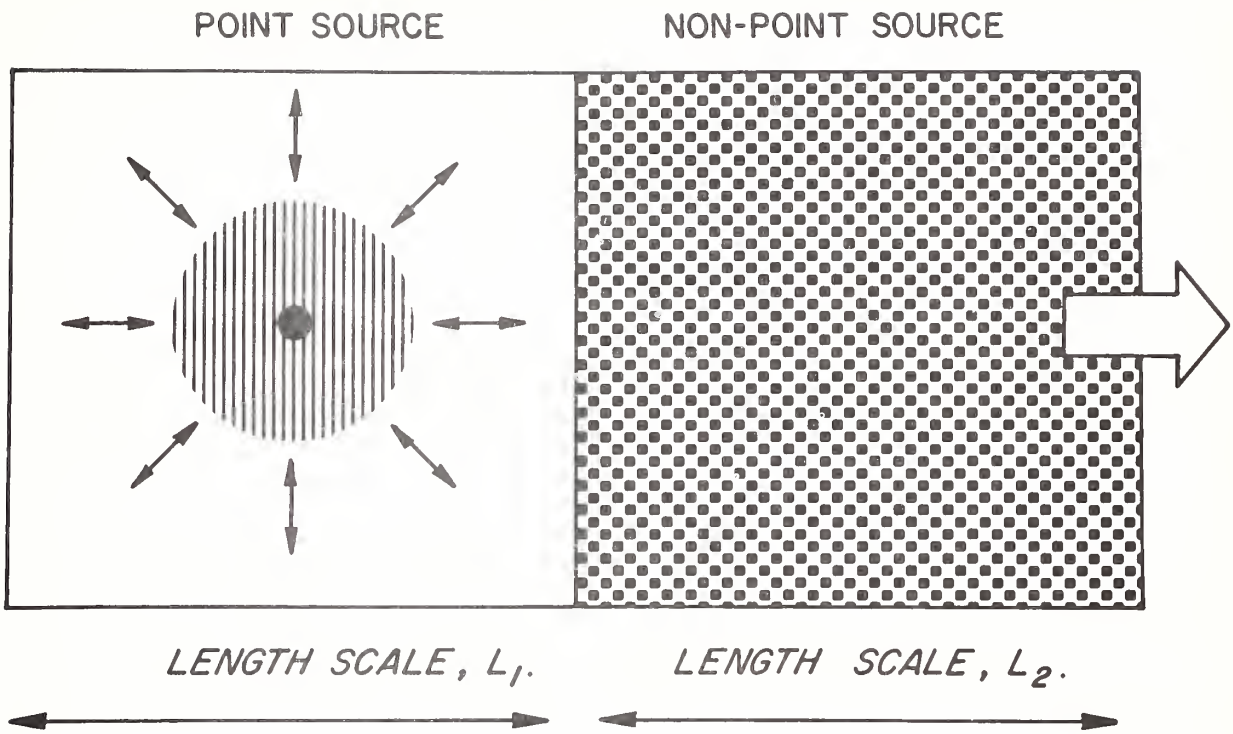


Figure 2.
Left - A point source of chemical, where the length scale L1 may be of the order 1 m to 10 m. Right - A non point source of chemical, where the length scale L2 may be of the order 10 m to 10 km.



Figure 3.
A non point source of nitrogen from pasture, grazed by sheep, abutting the Tiritea Stream, Palmerston North, New Zealand. At a finer scale the non point source can be seen to comprise a host of urination point sources. What would be an appropriate surface N concentration to use in a non point simulation model?

achieved by a process of iterative curve-fitting within the process subroutines. This way the overallmodel can be made to appear to work (Passioura 1973). Unfortunately it is often neither attempted, nor practicable to assess the veracity of the mechanistic subroutines in grandiose models. Ronald Reagan's "doveryai no proveryai" (trust, but verify) could equally apply to computer modeling. Passioura (1973) concluded that "... crop simulation is at present the art of the plausible. As such it is closer to metaphysics, than it is to science". More recently Monteith (1981) suggested "... perhaps we should declare a moratorium on the more sophisticated forms of modeling (quadrant II) until physiological work catches up (e.g. quadrant I) ... we need to know more about the way in which roots (i.e. quadrant IV) respond to soil physical conditions (i.e. quadrant III)" (parentheses added). Undoubtedly "... simulation seems to offer a good point of departure for theorists and experimentalists to begin journeying together" (Hillel, 1977), however a disparate pace of progress has developed. The experimentalist has been unable to match the speed of the computer. In a relatively recent treatise on the simulation of nitrogen behavior of soil-plant systems (Frissel and van Veen 1981) Addiscott et al. reviewed migration processes in soil (i.e. quadrants III & IV) and concluded that "... our mathematical skills still exceed our knowledge about the biological system being described and the quality of the input data available. Thus, care should be taken when evaluating output from nitrogen models based on minimal knowledge and inadequate input data".

IMPERATIVES

From my unashamedly-biased perspective, research imperatives are now outlined that might lead to better assessment of the effects of management on the quality of water passing through or across tracts of agricultural land. Not surprisingly these imperatives focus on attainment of a better understanding of the small-scale processes that operate within non point sources.

It is suggested that there is a need to focus more finely on the processes operating at the soil surface and in the rooted region immediately beneath. With the acknowledged bias of a soil-water physicist, it is considered essential that primacy be given the entry, root extraction, and redistribution of water in surface soil. Water is not only a solvent of many applied and resident soil chemicals, it is the vehicle by which such chemicals are transported into surface and groundwaters. This soil water emphasis is reasonable, for one of the more common forms of land management that creates water quality problems is irrigation, or more particularly unknowing or inadvertent over-irrigation (table 1). We need only be preoccupied with the soil surface, and the biologically-active, organic-matter rich horizons exploited by roots. Once soil water and its passenger chemicals have progressed beyond this zone there is little chance of natural recovery before they become groundwater. Furthermore the soil condition at the immediate surface is most affected by land management. But it is exactly here that water and other chemicals are prejudicially despatched to the plant roots, or consigned to either surface or ground waters. Also it will be more productive to research and experiment with surface soil for the simple reason that it is easier to get at.

Four processes that relate critically to issues of water quality are identified as being inadequately understood for the purposes of modeling the consequences of land management.

(1) The role of the aerial parts of plants, and crop architecture on controlling the pattern and flux density of water and chemical arrival at the soil surface.
(2) The effect of the physical condition of the immediate soil surface on the partitioning of surface water into either infiltration or local runoff, and the effect of surface-soil structure on the pattern of water and chemical entry.
(3) The ramification and exploitation of surface soil by roots, their dynamics and behavior, and especially the rate and pattern of water and chemical uptake.
(4) The redistribution of water within the root zone following infiltration.

A thread common to the first three imperatives is that soil physicists will need to broaden their perspective beyond the physical realm and address biological and chemical concerns that presently can only be described qualitatively. This nescience is neither the fault of the theoretician nor the experimentalist, but simply represents the fearful complexity of the maze of edaphic processes. The fourth imperative is entirely within the ambit of soil physics. It is just that our mathematics has not been as successful at describing the complicated, hysteretic, unsaturated process of redistribution, as it has been in resolving monotonic infiltration which possesses a simpler boundary condition. A major concern remains one of measurement. We cannot yet measure soil or plant properties anywhere near as easily as we can solve an equation or carry out a computer simulation.

Crop Architecture

When the local flux density of water (J_w, $m^3/m^2/s$) falling on the surface is less than the saturated hydraulic conductivity of the soil matrix ($K_s(\bullet)$, m/s), the pressure potential at the soil surface (ψ_0, m) will remain negative (Rubin 1966). The applied water will be absorbed directly into the unsaturated soil. For this third, or flux-type, boundary condition the spatial pattern of wetting will directly mimic the spatial pattern of application. However should J_w remain greater than $K_s(\bullet)$ then eventually $\psi_o \rightarrow 0$. When $\psi_o = 0$ for the first time t=tp, and incipient ponding is said to have occurred (White et al. 1982). This creation of a free-water film on the soil surface has two very important effects on the pattern of soil wetting.

When t>tp and $\psi_0 = 0$, then the surface film of water is free to enter any surface-vented macropores. These macropores often allow the rapid and far-reaching transport of water and chemical into and through the root zone. The pattern of wetting can be chaotic and the surface area of soil exposed to the invading solution may be quite small (White et al. 1986). A necessary condition for large macropores to operate in this manner is that ψ_0 is near-zero at their entrance. This can result simply from ponding of water on the soil surface (a first, or concentration-type boundary condition), or this condition is met when J_w remains sufficiently larger than $K_s(\bullet)$ long enough for t to exceed tp.

Secondly when the surface-soil matrix ponds, the ponded condition $\psi_0(A) = 0$ is achieved over some local region A, and the film of surface free water can move downhill by gravity flow. The efflux of free-water from A may well be accommodated by the additional surface area of sorption created by entry into the macropore system. Failing this, local runoff from A may, by downstream accumulation swell to generate a larger scale of runoff that might rapidly reach nearby surface water bodies. This interaction of surface-vented macroporosity and runoff was illustrated by Sharpley et al. (1979). They reported 1650 m^3/ha/yr of surface runoff from a 13° slope of pasture inhabited by surface-casting earthworms (Lumbricus and Allolobophora spp.). But they found runoff to double to 3210 m^3/ha/yr from similar pasture in which these worms were eliminated by repeated spraying with carbaryl. Total runoff loss of chemical was 0.91 and 1.8 kg-P/ha/yr, and 4.7 and 17.3 kg-N/ha/yr for the earthworm and carbaryl treatments respectively. The creation and maintenance of a surface free-water film is critical, for it prejudices the entry of water and chemical into the surface soil. The flux criterion for establishment of $\psi_0(A) = 0$ is that over A, J_w exceeds $K_s(\bullet)$ for a sufficient period of time. In this context we consider the role of vegetation in manipulating the spatial pattern of $J_w(A)$ during either natural rainfall or sprinkler irrigation. Ironically when studying infiltration, soil physicists often demonstrate their preoccupation with things 'physical' by first shaving the soil surface clear of any confounding vegetation (Clothier and White 1981)! However on agricultural land, water first falls on plants and by them is subsequently transmitted to the soil surface. The spatial pattern of the arrival of water at the soil surface can be strongly controlled by the architecture of the vegetation.

Of importance is the lens-like role that the vegetation may play in focusing the applied water. Locally, $J_w(A)$ may swell to a level well above the areal-average rate. Plants may act as

reverse-umbrellas whereby leaves or branches intercept falling water and direct it inwards to the stem or trunk. Thus at the soil surface surrounding the base of the plant the flux density of water may be large and greater than K_s•. So ponding ensues rapidly and is maintained over some adjacent area by stem flow. Plants create and sustain many of the large, continuous macropores that are exploited by surface free-water. A focal point of this biogenic macropore network is often at the soil surface, right at the base of the stem or trunk - exactly where the plant itself maintains ponding by focusing J_w. This anastomosis may encourage the rapid transmission of applied water directly and preferentially into the root zone. Obviously the architecture of the plant canopy and the extent of surface-vented macroporosity are crucial for directing water and chemical either into the soil or across the surface. This will vary greatly with tillage method, crop and soil type, land use and management.

Interception and stem flow have long been considered by hydrologists and foresters (Horton 1940, Rutter 1964), however scant attention has been given their role in soil wetting and root zone chemical transport. Agriculturalists have been remiss, with a few exceptions, in studying the effect of vegetation in directing water to the soil surface. Haynes (1940) reported that 20 percent of rain falling on corn or soybean arrived at the soil surface via the stem. Saffigna et al. (1976) found this to be 20-40 percent for ridged potatoes and observed very non uniform soil wetting as a result. Jury et al. (1976) discussed the implications of this non uniformity on fertilizer placement and preferential leaching. Kanchanasut and Scotter (1982) applied Rhodamine-B dye to pasture and after 40 days observed a far deeper penetration of dye under the crowns of the individual plant, as shown in figure 4. Infiltration patterns established by such fractional wetting are more likely to persist (Raats, 1973) than become more uniform (Philip 1983).

Water application methods, vegetative cover and soil structure can interact to produce quite different leaching patterns. Little consideration seems given to these critical, localized soil-surface processes in current models of water quality assessment. Before this can be achieved, better understanding of the mechanisms of water arrival at the soil surface will need to be obtained via experimentation and keen observation.

Surface-Soil Structure

The flux criterion for attainment of incipient ponding (ψ_0=0) is that the local flux density of water, J_w, be greater than the saturated hydraulic conductivity of the matrix, K_s•. This matrix conductivity is intended here to be the permeability of the body of the soil, not including non-capillary macropores with orifices greater than about 1 mm. This distinction defines a matrix-macropore dichotomy wherein the matrix exhibits the hydraulic properties expected of the soil on the basis of texture, and the macropore system reflects intimately the gross structural state of the soil. Soil physics theory (Buckingham 1907, Philip 1969) has been useful in describing the pattern of water and chemical movement through the matrix of uniform, isotropic soils. However theory has been less successful in dealing with the complexity of flow in macropores. The soil-surface macropore system, which only operates at potentials very close to saturation (say ψ_0>-30 mm) can result in rapid and far-reaching transport of water, chemical, and even particulate matter. The macropore system is often interconnected and chaotic, frequently being biogenic or pedogenic. Surface macroporosity is fragile and easily disturbed by soil management.

The saturated matrix conductivity (K_s•) of field soils tends to be of the order 10 mm/hr (White and Sully 1986, Watson and Luxmoore 1986, Moore et al. 1986). Hence rainstorms or irrigation that commonly have J_w greater than about 15-30 mm/hr will result in some localized ponding on the surface matrix, even without consideration of vegetation funneling. The presence of surface free-water means that the macropore system is a frequent and dominating mode of water entry and transport in surface soil during infiltration (Thomas and Phillips 1979). Unfortunately this renders inappropriate many of the hard-won gains of uniform soil physics theory, for such

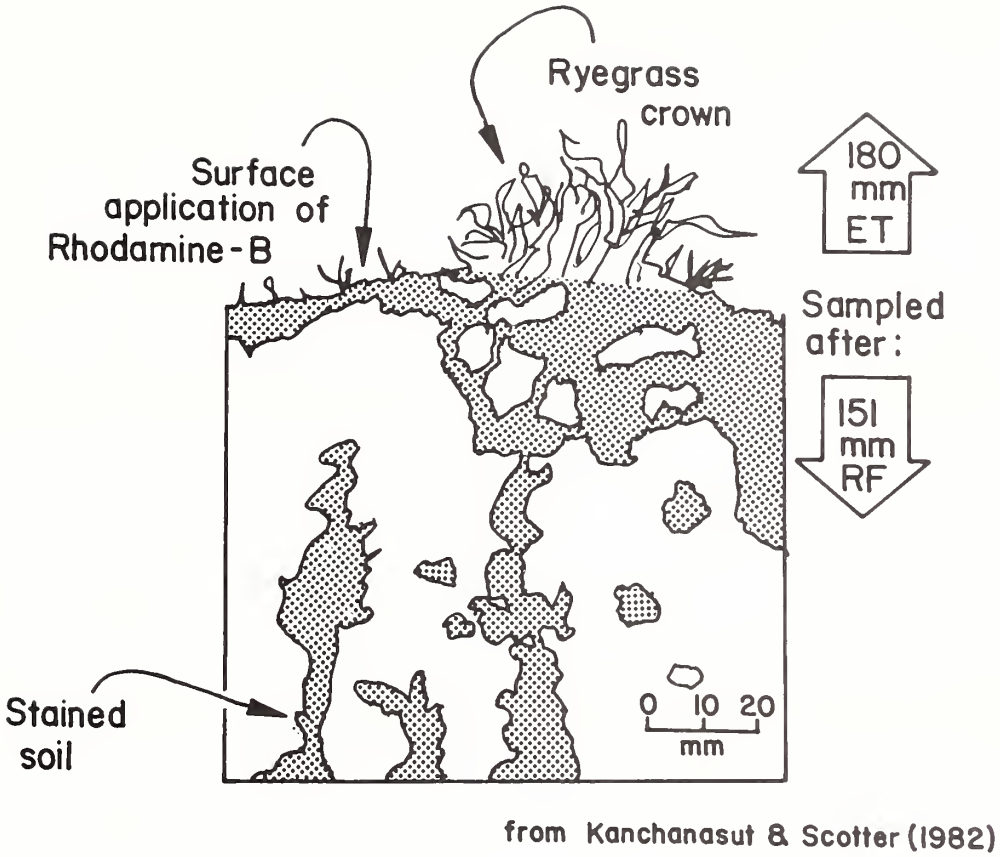


Figure 4.
Dye penetration under a ryegrass plant reveals an anastomosis at the surface between water flow from the stems, and macropore flow into the soil immediately under the crown (from Kanchanasut and Scotter 1982).

predictions apply only to the more sedate, unsaturated flow through the matrix which is commonly a second-order process. When simulating it is imperative to model the appropriate processes.

Drainage is commonly used as a management strategy to allow productive use of land bedeviled by excess water. One drainage system uses mole-plugs pulled through the soil at a depth of about 400 mm every 2 m perpendicular to the tile drains (Leeds-Harrison et al. 1982).

Theoretical solutions can be found for the uniform flow of free-water through soil to mole drains (Warrick and Kirkham 1969). Travel-time estimates can then be made to predict the breakthrough of surface-applied chemical in mole-drain effluent (Raats 1978, Jury 1975). The arrival time of any chemical will reflect the velocity field of water movement, as well as any retardation resulting from adsorption of chemical onto soil particles. However Scotter and Kanchanasut (1981) found that strongly-adsorbed phosphorus appeared in mole-drain effluent virtually at the same time as non-adsorbed chloride. Both arrived well before the time expected on the basis of a uniform flow field (figure 5). Kanchanasut et al. (1978) even reported urea and coliforms in mole-drains within 2 hours of spraying effluent on the surface at 10 mm/hr. Some quick digging with a spade reveals the biogenic macropore network responsible for this rapid and preferential transport, shown in figure 6. Surface ponding of water is considered to occur rapidly on the matrix. This allows the rapid movement of a free-water film to macropore vents around the disturbance created by the mole blade, which are then maintained by plants. Thus water is quickly transmitted through the soil and easily enters the mole drain. The invading solution might be exposed to only a very small

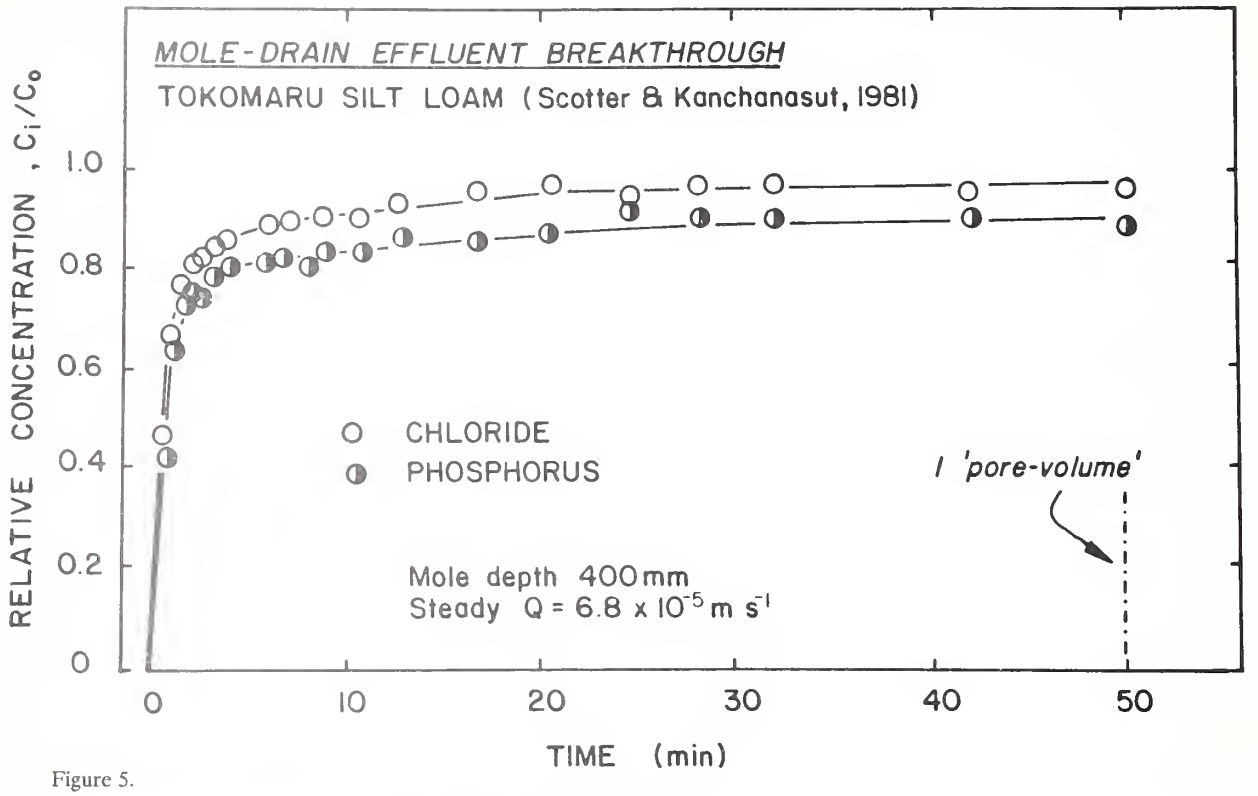


Figure 5.
The breakthrough of chloride and phosphorus in mole-drain effluent during steady discharge following application via an infiltrometer ring on pasture above the mole (from Scotter and Kanchanasut 1981).





Figure 6.
Left - The tap-root of a lucerne plant (alfalfa) following the crack made by the mole plow blade and directly entering the mole drain in Tokomaru silt loam. Above - A view up a mole drain showing the maze of fine lucerne roots penetrating the ceiling of the drain.

fraction of the soil volume. Expected (or hoped for) adsorption of chemical (Baes and Sharp 1983) need not occur.

Soil tillage and crop residue management can drastically alter the pattern of water entry into soil (Hamblin and Tennant 1981). Differences are likely in the hydraulic character of the matrix of mold-board plowed soil, and the same soil that has been minimum-tilled. Thus soil may have different times-to-ponding, depending upon tillage technique (Packer et al. 1984). Tillage also substantially modifies the surface-vented macroporosity and surface micro-relief, and this dictates the routing of the post-ponding film of free-water (Moore and Lawson 1979). Therefore the pattern of water entry into cultivated soil can depend considerably upon the method of tillage. Furthermore tillage can modify the profile of the soil's organic matter. The pattern of chemical entry and retardation will likewise be affected (Culley et al. 1987).

The rapid transit of water and chemical through surface soil is of concern from a water quality standpoint. Simulation and modeling via conventional theory is unlikely to be relevant (Reid and Parkinson 1984). Realization of the prime role played by macropores in the field has led to a burgeoning growth in preferential flow studies (Scotter 1978, Beven and Germann 1981, Bouma and Raats 1984). Unfortunately the mathematics we have at our disposal has not been as successful at describing chaotic macropore flow as it has been in resolving unsaturated matrix flow. Prospects are bright. Progress in understanding and description will follow field application of new measurement techniques, and development of appropriate theory (White 1988, Davidson 1987, Beven and Clarke 1986).

The qualitative matrix-macropore scheme of Dixon and Petersen (1971) has provided us with a useful model for gedanken simulations (figure 7). The need is to transfer these thoughts into quantitative realities. Free-water generated by surface ponding on the matrix can rapidly exploit macropores and be transported preferentially. Soil texture is a poor reckoner of hydraulic conductivity. Clothier and Heiler (1983) found that the saturated hydraulic conductivity of a well-structured clay loam to be 95.2 mm/hr; a poorly managed, surface-slaked silt loam was found much less permeable to free-water at 10.9 mm/hr. Simulation modelers require better information as to the effect of management and surface-soil structure on the hydraulic conductivity function K(U). There will be a need to establish more realistically the hydraulic conductivity function, especially at the wet end where the role of macropores is paramount. Here management can substantially modify the permeability by altering surface-venting and sub-surface connectedness (figure 8). Because of the proclivity of surface-venting earthworms for moisture (Edwards and Lofty 1977), Trout et al. (1987) found the infiltration rate of an Idahoan soil to increase 1.5 to 3 times during a 20-hr furrow irrigation.

New measurement techniques (Clothier and White 1981, Perroux and White 1988, Ankeny et al. 1988) coupled with versatile nonlinear parametric models of K(U) offer hope (Broadbridge and White 1988). These studies are beginning to reveal as inappropriate many of our currently-held views on water flow (Chong and Green 1983, Watson and Luxmoore 1986). Our preconceptions of the capillary length and time scales of field infiltration might be erroneous (White and Sully 1987).

Root Exploration and Exploitation

The aerial parts of plants often act to focus incident water towards a region of the soil surface where plant-engendered, surface-vented macropores can rapidly transmit water and chemical into the root zone. By this means plants appear to subvert and overcome the slow and retarded movement of water and chemicals that uniform soil physics theory would predict. Once in the root zone, plant uptake can proceed. Prediction of water and nutrient uptake by roots is currently described in most simulation models using some form of Gardner's (1960) seminal analysis of water movement to a stationary, line-sink 'root'. Ironically Gardner (1985) now feels that this has led us to a dead end! Breteler et al. (1981) concluded "... that the understanding of water

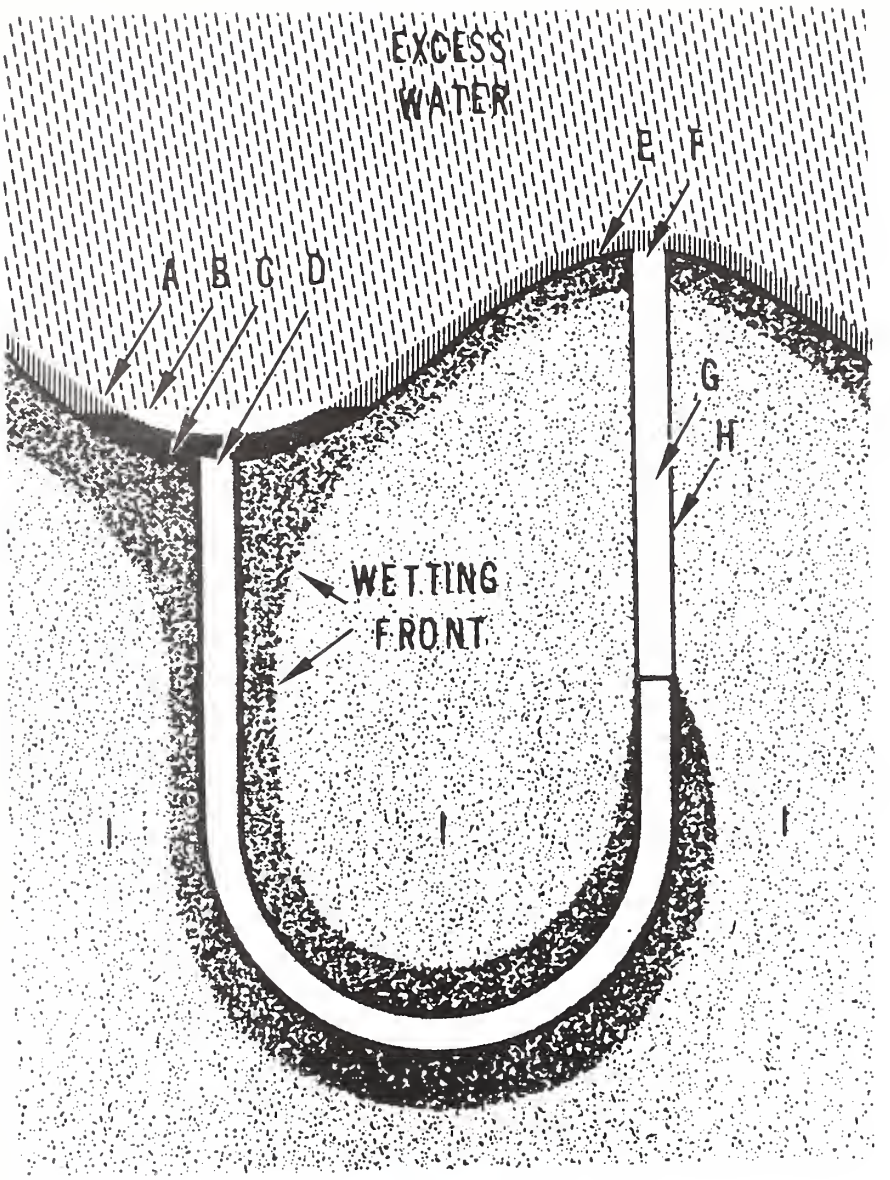


Figure 7.
The idealized matrix-macropore soil system of Dixon and Petersen (1971). Ponding on the flanks of the micro-catchment (A,E) creates a pond (B) in the micro-depression (C) so that free water enters the macropore (D,G) while air is exhausted (F) and water is absorbed into the unsaturated matrix (I).

transport in soils and plants is much greater than our understanding of root development, root function and ... nutrition ... major limitations to progress are the lack of technology for studying roots in the soil and the problems of validating complex models".

An imperative is better modeling of water and chemical interception by stationary and growing roots. In the first instance easier and more efficient field techniques for measuring root length density are required. Roots by virtue of being out of sight, have tended to be out of mind. Also root sampling has been difficult and tedious (Böhm 1979). But in prospect, results will arrive more easily from improved coring and washing devices (Smucker et al. 1982), as well as from non-destructive in situ techniques for recording rooting (Gregory 1979). These techniques will furnish data from which improved analytical descriptions of the topology of rooting can be

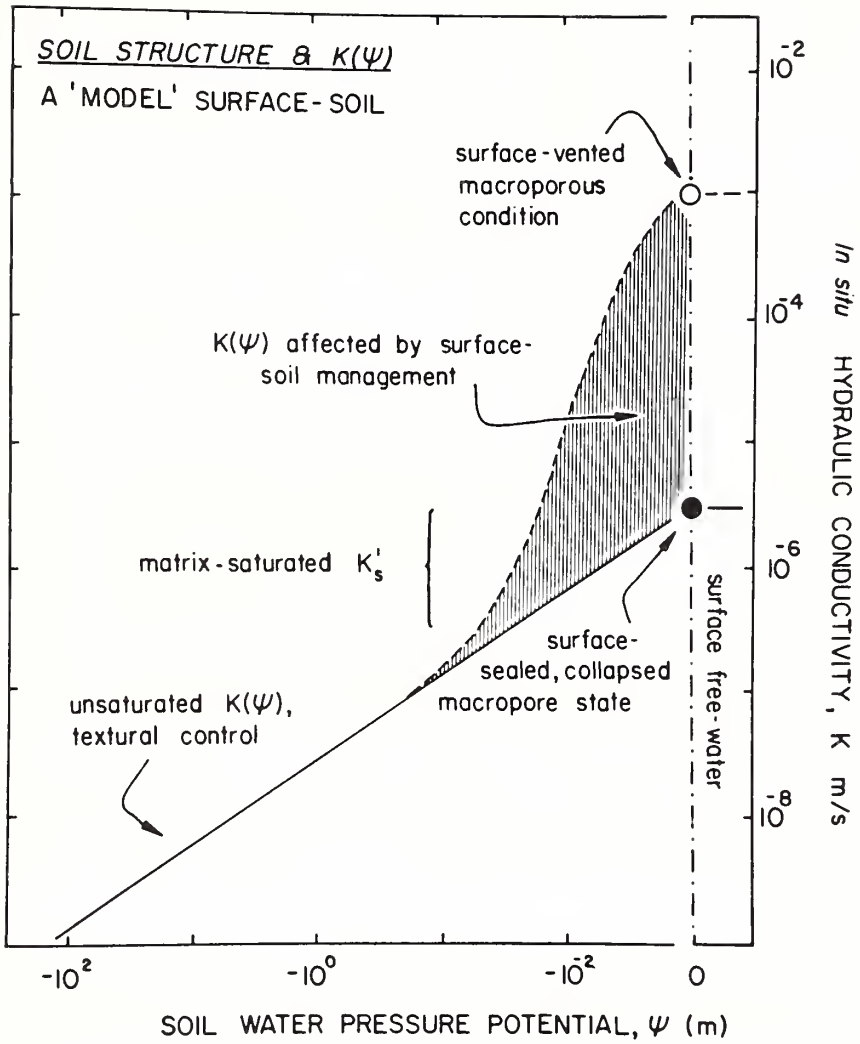


Figure 8.
A translation of the matrix-macropore notions embodied in figure 7 into a hydraulic conductivity function, K(ψ). The shaded area identifies the modification that soil management might impart upon the permeability. Also identified is the saturated conductivity of the matrix (K_s•), i.e. the soil minus those macropores that drain when Uo is less than about -30 mm.

developed (Gandar and Hughes 1988). Figure 9, taken from de Willigen and van Noordwijk (1987), shows the horizontal pattern of rooting found under a wheat crop tilled by four different devices, and the allocation of soil zones to the nearest root by Dirichlet tessellation. Management can affect ramification of soil by roots, especially given the proclivity of roots to grow along cracks and structural weaknesses (Hewitt and Dexter 1984). de Willigen and van Noordwijk (1987) analyzed the effect of the contrasting root distributions of figure 9 on the uptake of adsorbed chemical. They found that the rotadigger would allow the greatest unconstrained uptake.

Probably of greater practical concern than understanding the impact of root length density on water extraction and chemical uptake, is to assess more correctly the extent of rooting. A first-order requirement is prediction of the size and shape of the volume of soil explored by roots. It is the extent of the root zone that establishes the magnitude of the buffer against export of water and chemical to groundwater. Shallow root zones are more likely to be transparent to any invading soil water and its passenger chemicals. Currently a spade appears the best device for

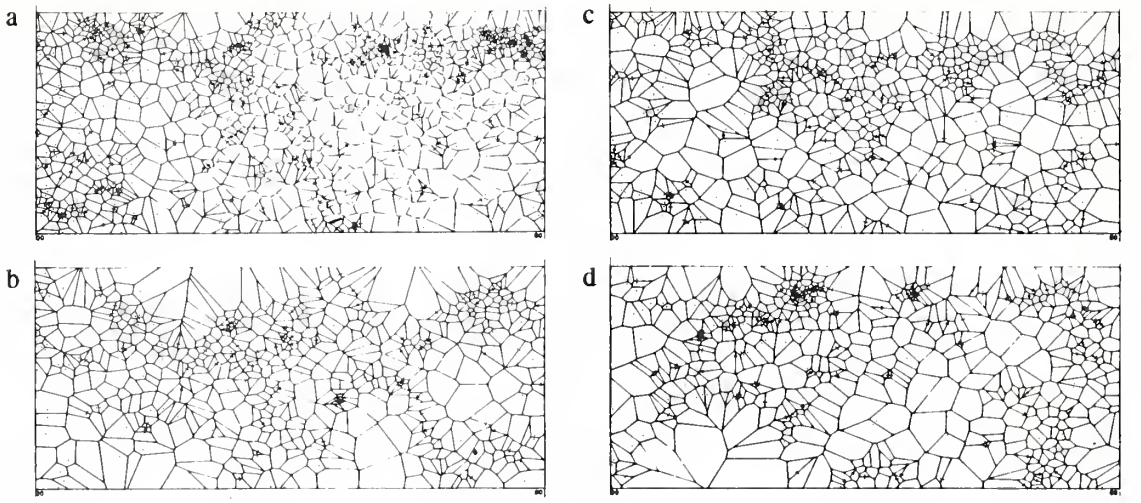


Figure 9.
Root maps under wheat (600x300 mm), and allocation of soil zones to the nearest root. Tillage was performed by (a) moldboard plow, (b) paraplow, (c) cultivator and (d) a rotadigger (from de Willigen and van Noordwijk 1987).

determining the extent of rooting. This is nevertheless a parameter critical for the success of sophisticated computer simulation models.

Monteith (1986) signposted the pre-eminent role of roots by suggesting that for crops we should "... model the supply system in terms of root performance, rather than the complex behavior of the stomatal valve". For annual crops he considered that upon reaching a maximum root-front penetration velocity of about 20-40 mm/day, plant water use is limited by the supply of water yielded by root penetration. An exception occurs when the soil behind the root front is rewetted by rain or irrigation: then plant water use becomes limited by demand. It will be interesting to see if simulation modelers can feasibly adopt this different approach. The increasing availability of root penetration data (e.g. Hansson and Andren 1987) will make the modeling task easier.

The problem of root penetration is more complicated for isolated plants, such as in forests, or with widely-spaced horticultural crops. Here there is a need to consider the three-dimensional penetration of roots. Kiwifruit vines are commonly planted in irrigated orchards at a 5x5 m spacing. Clothier et al. (1986) observed 'dry-spots' in the bare soil of the herbicide-strip around unirrigated vines. They attributed these spots to root-water extraction, and from vines of various ages they inferred a radial rate of growth in the root-front of 225 mm/yr. Snow (1987) monitored by neutron probe the changing soil water content around a kiwifruit vine covered at ground level to eliminate infiltration or evaporation. With no drainage the measured change in profile water storage, W(r), reflects only root-water extraction. So to depth z^* at distance r from the vine,

$$W(r) = \int_0^{z^*} \frac{\partial \theta(r,z)}{\partial t} \, dz \qquad [1]$$

As shown in figure 10 the radial pattern of W(r) reflects the radial pattern of rooting. The dry spot radius associated with the root front was observed at 1.5 m for this 8 yr old vine in the summer of 1985/86. It takes about 10 years for the vines to fill in completely. So in the interim the rate of vine water-use must be calculated using,

$$Q = 2\pi \int_0^{r^*} rW(r)dr \qquad [2]$$

which in the present case amounts to about 30 liters/day.

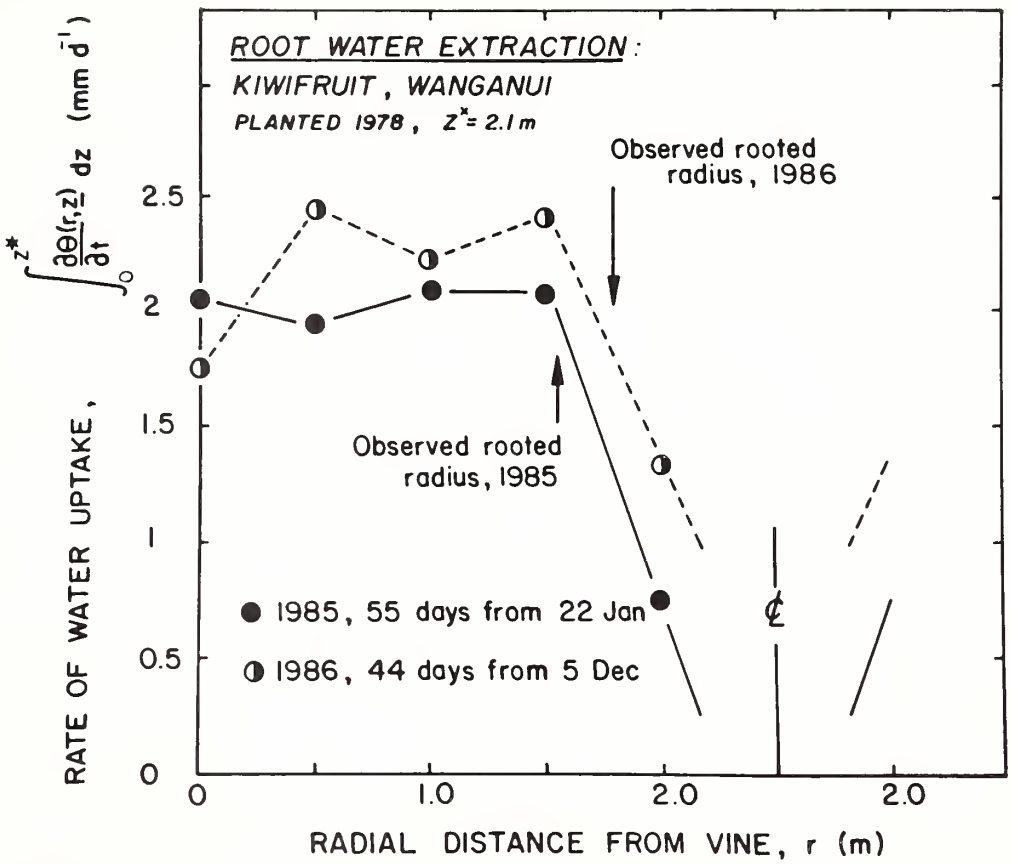


Figure 10.
The radial pattern of root-water extraction around a kiwifruit vine covered at ground level to prevent infiltration and soil evaporation (from Snow 1987). Also identified are the limits of rooting inferred from observations of surface dry spots.

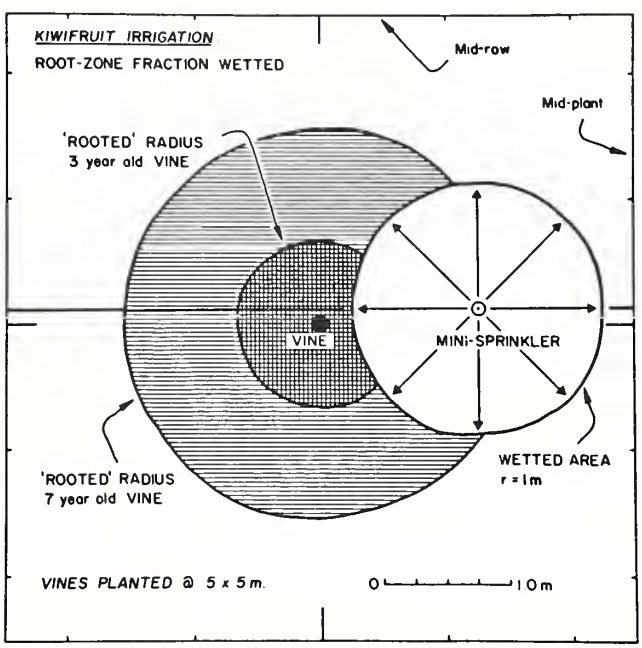


Figure 11.
A typical minisprinkler-irrigated kiwifruit vine. The minisprinkler is presumed to have a radial throw of 1 m, and the rooted fraction of the plan area is shown for 3 and 7 year-old vines.

The expanding rootzone has implications for irrigation efficiency, orchard management and water quality. If irrigation and dissolved chemicals are applied to young vines via dripper or minisprinkler, (e.g. Gerst and Albasel 1984, Clothier and Sauer 1988) it may be that some large fraction fails to intercept the roots (figure 11). Deleterious environmental consequences must ensue if this applied water and chemical are progressively leached through rootless soil. As the vine grows, the intercepted fraction should rise so that an increasing proportion of the applied chemical is made available in the root zone (quadrant III).

A prime requirement for better simulation of the water and chemical balance of crops is additional information and knowledge about the exploration of soil by plant roots. Secondly it will be important to gain greater understanding of the manner by which plant roots exploit this explored soil. This will permit better modeling of the attainment by roots of water and chemical from surface soil (quadrant IV).

Redistribution

Once water has infiltrated and during root-water uptake, redistribution of water occurs within the soil in response to gravity and spatial gradients of soil-water pressure potential. It is the balance between root-water extraction and redistribution that establishes the fraction of the invading water and chemical that becomes available in the root zone (quadrant III). The complementary portion ends up draining beyond the root zone to groundwater.

Infiltration into uniform soil has been successfully simulated by either analytical (Philip 1969), or numerical (Campbell 1985) solution of the soil water flow equation,

$$\frac{\partial \theta}{\partial t} = \left[K(\psi) \frac{\partial \psi}{\partial z} \right] + \frac{\partial K}{\partial z}, \qquad [3]$$

subject either to surface ponding,

$$\theta = \theta_o, \; z = 0, \; t \geq 0 \qquad [4]$$

or some prescribed surface flux, V_o,

$$K \left(\frac{\partial \psi}{\partial z} - 1 \right) = V_o, \; z = 0, \; t \geq 0 \qquad [5]$$

Given $K(\psi)$ it is possible to simulate infiltration into uniform soil (figure 12). But it is difficult to measure the unsaturated hydraulic conductivity function, $K(\psi)$, even in the laboratory. Thus the form of $K(\psi)$ is frequently obtained by a 'fix' at saturation with the measured K_s. Then the unsaturated portion is inferred from the soil's texture, or commonly deduced by some means from the shape of the water retention curve, $\psi(\theta)$, (Childs and Collis-George 1950). For example

$$K(\theta) = K_s \left(\frac{\theta}{\theta_s} \right)^m, \qquad [6]$$

where the Brooks and Corey (1966) exponent, $m=2b+3$, comes from

$$\psi(\theta) = \psi_e \left(\frac{\theta}{\theta_s} \right)^{-b}. \qquad [7]$$

Ponded, or near-ponded infiltration is controlled to a large extent by the physical properties of the soil at the wet end of the scale, so use of 'saturation-matching' of K often allows successful simulation. When the wetting of infiltration ceases, the surface boundary condition is replaced by one of no flux. Redistribution then proceeds with unsaturated draining of the soil near the surface, and further unsaturated wetting of the soil at depth. Figure 12 shows two days of redistribution in a laboratory column of Manawatu silt loam following one hour of flux infiltration. The mixture of wetting and draining may mean that hysteresis should be considered. A successful simulation of redistribution using Campbell's (1985) program is shown in figure 12 for this repacked soil (Scotter et al. 1988). The simulation used the measured $K(\theta)$, and the appropriate wetting or scanning portion of $U(\theta)$. However if the 'saturation-matched' $K(\theta)$ function is used, the simulation erroneously shows water redistributing to over one meter deep. This failure is not due to ignorance of hysteresis, but that the simple parameterization of $K(\theta)$ grossly overpredicts the conductivity at the lower water contents relevant for unsaturated redistribution. The saturation matching of most simple $K(\theta)$ models, while appropriate for infiltration and near-saturated flow, turns out to be inadequate for the unsaturated redistribution that may operate for many days following infiltration. For this uniform, repacked laboratory soil Scotter et al. (1988) did however manage to simulate redistribution successfully, but only when

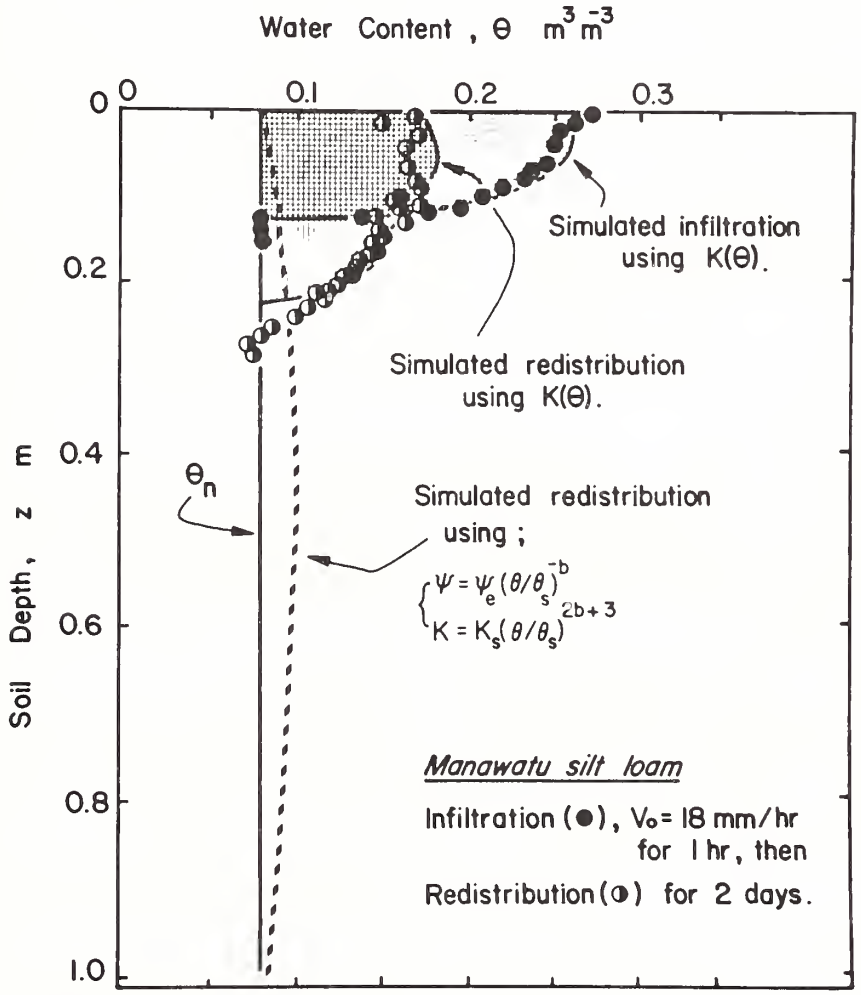


Figure 12.
Flux infiltration and redistribution in a vertical column of repacked Manawatu silt loam. A simulation of infiltration and hysteretic redistribution (——) using the measured $K(\psi)$ is shown. Simulation of redistribution with a 'saturation-matched' $K(\psi)$ is also shown (---) (from Scotter et al. 1988).

$K(\theta)$ was matched to a measured, unsaturated value of K in the appropriate range of water contents. This emphasizes the critical need for apt measurement and rational parameterization in simulation models of water and chemical redistribution in soil.

These problems encountered in the laboratory pale in comparison with the complexity met in the field. Natural layering in the soil profile can be confounding, and the redistribution process may compete against root-water uptake. These complications reinforce the requirement for appropriate measurement and description. Infiltration, redistribution and root-water uptake were measured in permeable sandy soil surrounding an irrigated apple tree covered at ground-level to eliminate extraneous evaporation (figure 13). Although the initial wetting around the tree nominally amounted to about 60 mm, surface ponding and lateral movement of surface free-water meant that the region surrounding this access tube received 109.3 mm. Neutron probing over the next 21 days, corroborated by heat-pulse measurement of sap-flow (Green and Clothier 1988), revealed 52.7 mm of root-water extraction from the layers within the nearly 2 m deep root zone. Root extraction of soil water was greatest near the soil surface, but in far greater proportion than the root length density data would simply suggest. Drainage into the coarse gravel layer beyond the root zone at 2 m accounted for about half the applied water, with obvious detriment to the groundwater and the immediately neighboring stream. Simulation of the field-water economy and chemical balance will require additional understanding of the role of root length density on plant uptake, as well as better measurement and parameterization of the soil's hydraulic properties.

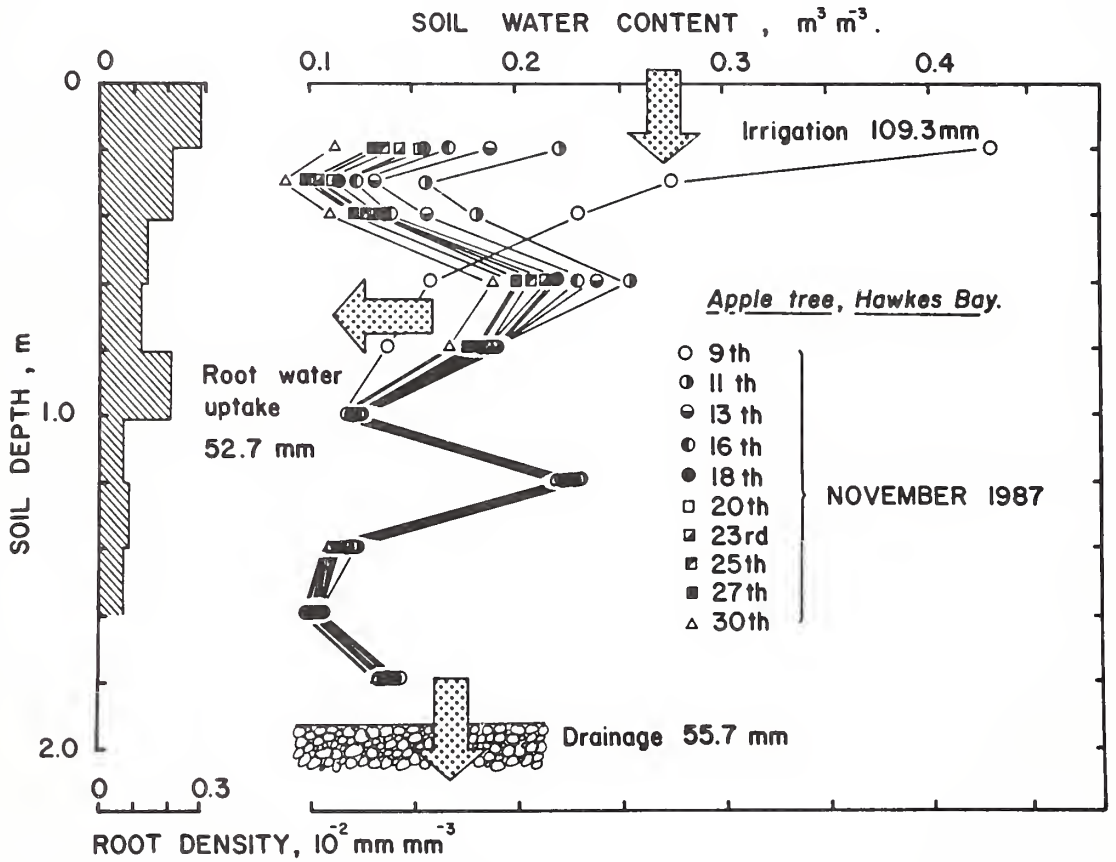


Figure 13.
The pattern of wetting, redistribution, root uptake and drainage measured in soil following irrigation of an apple tree. The neutron probe access tube was located 1.5 m from the tree. The 21-day water balance is given. The root length data were obtained by coring around nearby trees (from Rahardjo 1988, and K.A. Hughes and M. Peipi pers. comm.).

CONCLUSION

Prior to more sophisticated computer modeling of soil water flow and chemical transport to assess the impact of management on water quality, we need better measurements and understanding of the mechanisms operating at the point. These small-scale processes when integrated comprise the non point source: or as Yogi Berra might well say " The issue about non point sources depends on your point of view"!!! Caldwell and Russell (1987) stress the need for process-level modeling, for "... traditional field experimentation is inadequate for evaluating the (impact of) management. The number of possible management options ... is too great. There are also strong management by environment interactions".

Thankfully, measurement techniques and analytical procedures are developing rapidly. New infiltrometers have recently been designed that allow discrimination between macropore (Reynolds et al. 1985) and matrix flow (Bouma and Denning 1972, Dixon 1975, Dirksen 1975, Clothier and White 1981, Chong and Green 1983, Perroux and White 1988, Ankeny et al. 1988). Commensurate theory is also being developed to provide insight into the relative transport roles of macropore dimensions and matrix properties (Scotter 1978, Edwards et al. 1979, Beven and Germann 1982, Davidson 1987, Beven and Clarke 1986, Smettem 1986). Better parameterization of the unsaturated flow properties of field soil is being achieved (Broadbridge and White, 1988) and current dogma questioned (White and Sully 1987, Clothier et al. 1981). Also new techniques are giving a clearer picture of the biological environment of the soil. Time domain reflectometry (Topp and Davis 1981) can provide a clearer view of water movement very near the soil surface, and tomographic procedures are being applied to root-water extraction measurements (Hainsworth and Aylmore 1986). Subterranean periscopes, rhizotrons and automatic root-washing facilities are allowing better resolution of root growth and behavior. Analytical description of the pattern of rooting and its consequences is improving (de Willigen and van Noordwijk 1987).

In the first instance these new techniques and their results are tending to expose grave inadequacies in current simulation models, which tend to be based on less recent, poorly verified and probably out of date concepts and formulations (Ferrari 1965). When simulators are stimulated to tackle the new issues now being observed, their models will more faithfully represent real phenomena. Progress will come from keener observation and critical thinking, we should not expect "Deus ex machina".

REFERENCES

Addiscott, T.M., J.M. Davidson, K. Harrison, P.A. and others. 1981. Migration processes in soils. In M.J. Frissel and J.A. van Veen eds. Simulation of nitrogen behavior of soil-plant systems. 277 pp., Pudoc,Wageningen.

Andreski, S. 1972. Social Sciences as Sorcery Deutch (London).

Ankeny, M.D., T.C. Kaspar and R. Horton. 1988. A design for an automated tension infiltrometer. Agronomy Abstracts S-1:156.

Baes, C.F., and R.D. Sharp. 1983. A proposal for estimation of soil leaching and leaching constants for use in assessment models. J. Environ. Qual. 12:17-28.

Ball, P.R., and J.C. Ryden. 1984. Nitrogen relationships in intensively managed temperate grasslands. Plant and Soil 76:23-33.

Ball, P.R., D.R. Keeney, P.W. Theobald, and P. Nes. 1979. Nitrogen balance in urine-affected areas of a New Zealand pasture. Agron. J. 71:309-314.

Beven, K.J., and R.J. Clarke. 1986. On the variation of infiltration into a homogeneous soil matrix containing a population of macropores. Water Resour. Res. 22:383-388.

Beven, K.J., and P.F. German., 1982. Macropores and water flow in soils. Water Resour. Res. 18:1311-1325.

Böhm, W. 1979. Methods of studying root systems. Springer-Verlag, New York.

Bouma, J., and J.L. Denning. 1974. A comparison of hydraulic conductivities calculated with morphometric and physical methods. Soil Sci. Soc. Am. Proc. 38:124-127.

Bouma, J., and P.A.C. Raats eds. 1984. Proceedings of the ISSS Symposium on Water and Solute Movement in Heavy Clay Soil. ILRI pub. 37. Wageningen.

Breteler, H., D.J. Greenwood, I. Petterson, and others. 1981. Soil-plant relations. In M.J. Frissel and J.A. van Veen eds. Simulation of nitrogen behavior of soil-plant systems. 277 pp., Pudoc, Wageningen.

Broadbridge, P., and I. White. 1988. Constant-rate rainfall infiltration: A versatile nonlinear model. I. Analytic solution. Water Resour. Res. 24:145-155.

Brooks, R.H., and A.T. Corey. 1966. Properties of porous media affecting fluid flow. J. Irrig. Drainage Div., ASCE Proc. 72:61-68.

Buckingham, E. 1907. Studies in the movement of soil moisture. USDA, Bureau of Soils, Bull. 38.

Burden, R.J. 1982. Nitrate contamination of New Zealand aquifers: A review. N.Z. J. Sci. 25:205-220.

Caldwell, R.M. and J.T. Russell, 1987. The necessity of process-level modeling in intercrop research. Agronomy Abstracts A-3:11.

Campbell, G.S. 1985. Soil physics with BASIC. Developments in Soil Science 14, pp 150 Elsevier, N.Y.

Childs, E.C., and N. Collis-George. 1950. The permeability of porous materials. Proc. R. Soc. London, A. 201:392-405.

Chong, S.K., and R.E. Green,.1983. Sorptivity measurement and its application. In Proc. Nat. Conf. Adv. Infiltration, Chicago, ASAE Pub. 11-83, pp 82-91.

Clothier, B.E. and T.D. Heiler, 1983. Infiltration during sprinkler irrigation: theory and field results. In Proc. Nat. Conf. Adv. Infiltration, Chicago. ASAE Pub. 11-83:275-284.

Clothier, B.E., and D.R. Scotter. 1982. Constant-flux infiltration from a hemispherical cavity. Soil Sci. Soc. Am. J. 46:696-700.

Clothier, B.E., and I. White. 1981. Measurement of sorptivity and soil water diffusivity in the field. Soil Sci. Soc. Am. J. 45:241-245.

Clothier, B.E., J.H. Knight, and I. White. 1981. Burger's equation: application to field constant-flux infiltration. Soil Sci. 132:255-261.

Clothier, B.E., V.O. Snow, S.R. Green, and others. 1986. Water economy of kiwifruit vines. New Zealand Kiwifruit Authority Spec. Pub. 1:16-19.

Clothier, B.E. and T.J. Sauer, 1988. Nitrogen transport during drip fertigation with urea. Soil Sci. Soc. Am. J. 52:345-349.

Culley, J.L.B., W.E. Larson, R.R. Allmaras, and M.J. Shaffer. 1987. Soil water regimes of a Typic Haplaquoll under conventional and no-tillage. Soil Sci. Soc. Am. J. 51:1604-1610.

Davidson, M.R. 1987. Asymptotic infiltration into a soil which contains cracks or holes but whose surface is otherwise impermeable. Transp. Porous Media 2:165-176.

Dirksen, C. 1975. Determination of soil water diffusivity sorptivity measurement. Soil Sci. Soc. Am. Proc. 39:22-27.

Dixon, R.M. 1975. Design and use of closed-top infiltrometers. Soil Sci. Soc. Am. Proc. 39:755-763.

Dixon, R.M. and A.E. Peterson. 1977. Water infiltration control: A channel system concept. Soil Sci. Soc. Am. Proc. 35:968-973.

de Willigen, P., and M. van Noordwijk. 1987. Roots, plant production and nutrient use efficiency. Ph.D. thesis Agricultural University, Wageningen, The Netherlands, 282 pp.

de Wit, C.T. 1953. A physical theory of fertilizer placement. Ph.D. thesis, Agricultural University, Wageningen. 71 pp.

Edwards, C.A. and J.R. Lofty. 1977. Biology of earthworms. pp 333, Halsted Press, New York.

Edwards, W.M., R.R. van der Ploeg, and W. Ehlers. 1979. Numerical study of the effects on noncapillary-sized pores upon infiltration. Soil Sci. Soc. Am. J. 45:851-856.

Ferrari, Th.J. 1965. Models and their testing: considerations on the methodology of agricultural research. Neth. J. Agric. Sci. 13:366-377.

Field, T.R.O., P.W. Theobald, P.R. Ball, and B.E. Clothier. 1985. Leaching losses of nitrate from cattle urine applied to a lysimeter. Proc. Agron. Soc. N.Z. 15:137-141.

Frissel, M.J. and J.A. van Veen eds. 1981. Simulation of nitrogen behavior of soil plant systems. pp 277, Pudoc, Wageningen.

Gandar, P.W., and P.R. Ball. 1982. Nitrogen balances in New Zealand ecosystems. DSIR, Palm. Nth, pp 262.

Gandar, P.W., and K.A. Hughes. 1988. Kiwifruit root systems I. Root length densities. N.Z. J. Exp. Agric. (in press).

Gardner, W.R. 1960. Dynamic aspects of water availability to plants. Soil Sci. 89:63-73.

Gardner, W.R. 1985. Citation Classic. Current Contents Agric., Biol. Environ. Sci. 16:20.

Garfield, E. 1986. Hazardous Waste. Part I. The poisoning of our Plant. Current Contents 34:3-9.

Gerst, Z., and N. Albasel. 1984. Field distribution of pesticides applied via a drip irrigation system. Irrig. Sci. 5:181-193.

Gregory, P.J. 1979. A periscope method for observing root growth and distribution in field soil. J. Exp. Bot. 30:205-214.

Hainsworth, J.M., and L.A.G. Aylmore. 1986. Water extraction by single plant roots. Soil Sci. Soc. Am. J. 50:841-848.

Hamblin, A.P., and D. Tennant. 1981. The influence of tillage on soil-water behavior. Soil Sci. 132:233-239.

Hansson, A-C., and 0. Andren. 1987. Root dynamics in barley, lucerne, and meadow fescue investigated with a mini-rhizotron technique. Plant and Soil 103:33-38.

Haynes, J.L. 1940. Ground rainfall under vegetative canopy of crops. J. Am. Soc. Agron. 32:176-184.

Hewitt, J.S. and A.R. Dexter. 1984. The behavior of roots encountering cracks in soil. Plant and Soil 79: 11-28.

Hillel, D. 1977. Computer simulation of soil water dynamics: A compendium of recent work. Ottawa, IDRC, 214 pp.

Horton, R.E. 1940. Rainfall interception. Mon. Weather Rev. 47:603-623.

Hubbard, R.K., G.J. Gascho, J.E. Hook, and W.G. Knisel. 1986. Nitrate movement into shallow ground water through a coastal plain sand. Trans. ASAE. 29:1564-1571.

Jury, W.A. 1975. Solute travel-time estimates for tile-drained fields, II. Application to experimental studies. Soil Sci. Soc. Amer. Proc. 39:1020-1024.

Jury, W.A., W.R. Gardner, P.G. Saffigna, and C.B. Tanner. 1976. Model for predicting simultaneous movement of nitrate and water through a loamy soil. Soil Sci. 122:36-43.

Kanchanasut, P., and D.R. Scotter. 1982. Leaching patterns in soil under pasture and crop. Aust. J. Soil Res. 20:193-202.

Kanchanasut, P., D.R. Scotter, and R.W. Tillman. 1978. Preferential solute movement through larger soil voids. II. Experiments with saturated soil. Aust. J. Soil Res. 16:269-276.

Kirkham, D., and W.L. Powers, 1972. Advanced Soil Physics Wiley-Interscience, pp 534, New York.

Leeds-Harrison, P., G. Spoor, and R.J. Godwin. 1982. Water flow to mole drains. J. Agric. Eng. Res. 27:81-91.

Mandelbrot, B.B. 1982. The fractal geometry of Nature. W.H. Freeman.

Monteith, J.L. 1965. Evaporation and environment. In G.E. Fogg ed. The State and Movement of Water in Living Organisms, Symp. Soc. Exp. Bot. 19:205-234. Academic Press, N.Y.

Monteith, J.L. 1981. Epilogue: Themes and variations. Plant and Soil 58:305-309.

Monteith, J.L. 1986. How do crops manipulate water supply and demand. Phil. Trans. R. Soc. Lond A 316:245-259.

Moore, I.D., and C.L. Larson. 1979. Estimating micro-relief surface storage from point data. Trans. ASAE 22:1073-1077.

Moore, I.D., G.J. Burch, and P.J. Wallbrink. 1986. Preferential flow and hydraulic conductivity of forest soils. Soil Sci. Soc. Am. J. 50:876-881.

Nye, P.H., and P.B. Tinker. 1977. Solute movement in the soil-root system. Studies in Ecology 4, Blackwell, Oxford, pp. 342.

O.E.C.D. 1986. Water pollution by fertilizers and pesticides. Org. Econ. and Cult. Devel., Paris, pp. 144.

Packer, I.J., G.J. Hamilton, and I. White. 1984. Tillage practices to conserve soil and improve soil conditions. J. Soil Cons. N.S.W. 40:78-87.

Passioura, J.B. 1973. Sense and nonsense in crop simulation. J. Aust. Instit. Agric. Sci. 39:181-183.

Perroux, K.M., and I. White. 1988. Designs for disc permeameters. Soil Sci. Soc. Am. J. 52:1205-1215.

Philip, J.R. 1969. Theory of infiltration. Adv. Hydrosci. 5:215-296.

Philip, J.R. 1975. Soil-water physics and hydrologic systems. In Computer Simulation of Water Resources Systems, pp 85-102, North-Holland, Amsterdam.

Philip, J.R. 1983. Infiltration in one, two, and three dimensions. In Advances in Infiltration, pp 1-13 ASAE pub 11-83, Michigan.

Post, H.R. 1974. Against Idealogies (Inaugural lecture), Chelsea College, London.

Raats, P.A.C. 1973. Unstable wetting fronts in uniform and non-uniform soils. Soil Sci. Soc. Am. Proc. 37:681-685.

Raats, P.A.C. 1978. Convective transport of solutes by steady flows. I. General theory. Agric. Water Manage. 1:201-218.

Rahardjo, P. 1988. Soil water extraction around an apple tree. Dip. Agr. Sci. dissertation, pp 118. Massey University.

Reid, I., and R.J. Parkinson. 1984. Seasonal changes in soil-water redistribution processes affecting drain flow. In J. Bouma and P.A.C. Raats eds. Proc. ISSS symp. on water solute movement in heavy clay soils. ILRI publ. 37:156-160.

Reynolds, W.D., D.E. Elrick, and B.E. Clothier. 1985. The constant head well permeameter: Effect of unsaturated flow. Soil Sci. 139:172-180.

Rijtema, P. 1987. Nitrate load and water management of agricultural land. In Agricultural Water Management, A.L.M. van Wijk and J. Wesseling eds., pp.303-313, A.A. Balkema, Rotterdam.

Rubin, J. 1966. Theory of rainfall uptake by soils initially drier than their field capacity and its applications. Water Resour. Res. 2:739-749.

Rutherford, J.C., R.B. Williamson, and A.B. Cooper. 1987. Nitrogen phosphorus and oxygen dynamics in rivers. In Inland waters of New Zealand A.B. Viner ed. SIPC, DSIR Bull 241 Wellington pp.139-165.

Rutter, A.J. 1964. Studies in the water relations of Pinus sylvestris in plantation conditions. II. The annual cycle of soil moisture change and derived estimates of evaporation. J. Appl. Ecol. 1:29-44.

Saffigna, P.G., L.B. Tanner, and D.R. Keeney. 1976. Non-uniform infiltration under potato canopies caused by interception, stem flow, and hilling. Agron. J. 68:337-342.

Scotter, D.R. 1978. Preferential solute movement through larger soil voids. I. Some computations using simple theory. Aust. J. Soil Res. 16:257-267.

Scotter, D.R., and P. Kanchanasut. 1981. Anion movement in a soil under pasture. Aust. J. Soil Res. 19:299-307.

Scotter, D.R., B.E. Clothier, and T.J. Sauer. 1988. A critical assessment of the role of measured hydraulic properties in the simulation of absorption, infiltration and redistribution of soil water. Agric. Water Manage. 13:73-86 .

Sharpley, A.N., J.K. Syers, and J.A. Springett. 1979. Effect of surface-casting earthworms on the transport of phosphorus and nitrogen in surface runoff from pasture. Soil Biol. Biochem. 11:459-462.

Smucker, A.J.M., S.L. McBurney, and A.K. Srivastra. 1982. Quantitative separation of roots from compacted soil profiles by the hydropneumatic elutriation method. Agron. J. 74:500-503.

Smettem, K.R.J. 1986. Analysis of water flow from cylindrical macropores. Soil Sci. Soc. Am. J. 50:1139-1142.

Snow, V.O. 1987. The pattern of soil water extraction by individual kiwifruit vines. M. Ag. Sci. thesis, Massey University, pp. 107.

Sun, M. 1986. Groundwater ills: Many diagnoses, few remedies. Science 232:1490-1493.

Thomas, G.W., and R.E. Phillips. 1979. Consequences of water movement in macropores. J. Environ. Qual. 8:149-152.

Topp, G.C., and J.L. Davis. 1981. Detecting infiltration of water through soil cracks by time domain reflectometry. Geoderma 26:13-23.

Trout, T.J., W.D. Kemper, and G.S. Johnson. 1987. Earthworms cause furrow infiltration increase. Proc. Int. Conf. Infilt. Develop and Applic., Yu-Si Fok ed., Water Resour. Res. Centre, Honolulu, pp. 398-406.

Truesdell, C. 1984. The computer: ruin of science and threat to mankind. In An Idiot's Fugitive Essays on Science, pp.654, Springer-Verlag, New York.

Watson, K.W., and R.J. Luxmoore 1986. Estimating macroporosity in a forest watershed by use of a tension infiltrometer. Soil Sci. Soc. Am. J. 50:578-582.

Warrick, A.W., and D. Kirkham. 1969. Two dimensional seepage of ponded water to full ditch drains. Water Resour. Res. 5:685-693.

White, I. 1988. Measurement of soil physical properties in the field. In Flow and Transport in the Natural Environment, W.L. Steffen et al. eds., Springer-Verlag, Heidelberg.

White, I., B.E. Clothier, and D.E. Smiles. 1982. Pre-ponding constant-rate rainfall infiltration. In Modeling Components of Hydrologic Cycle. Water Resour. Publications, Colorado. 127-148.

White, I., and M.J. Sully. 1987. Macroscopic and microscopic capillary length and time scales from field infiltration. Water Resour. Res. 23:1514-1522.

White, R.E., J.S. Dyson, R.A. Haigh, and others. 1986. A transfer function model of solute transport through soil. 2. Illustrative Applications. Water Resour. Res. 22:248-254.

MANAGEMENT IMPROVEMENTS IN WATER QUALITY MODELS

Donn G. DeCoursey[1], James S. Schepers[2]

ABSTRACT

Many worldwide problems such as global climate change and water pollution call for solutions that are highly dependent on use of analytical models. These models are needed in research, for regulation, and for evaluation or assessment by action agencies. The use of models is motivated by the need for understanding, development of integrated management systems, identification of research priorities, process characterization and prediction. Numerous models have been developed, but many of these have failed. Those that failed did so because they could not be validated, were poorly planned, utilized poor software structure, or did not adequately simulate the desired processes. Features of water quality models that need emphasis include the dynamics of evapotranspiration, crop growth, nutrient cycling and soil water movement; the impact of tillage, freeze/thaw cycles, and biological processes on soil physical properties; recognition and inclusion of probabilistic features; inclusion of data bases; and development of more comprehensive research data. Future model development should evolve along the lines of a generic structure that is easier to verify, validate, maintain, interface with other models, and change. In the future, models will need to be more sophisticated, applicable to user's needs, and responsive to innovative management schemes that address climatic and environmental concerns.

INTRODUCTION

In the last few years, two issues have become overwhelming environmental concerns. These are the effects of global climate change and deteriorating water quality. Deteriorating water quality and its impact on quality of life was partially responsible for the U.S. Department of Agriculture's (USDA) Office of International Cooperation and Development (OICD) calling the International Symposium on Water Quality Modeling of Agricultural Non-Point Sources. In this paper, we address the needs of managers and policy makers as they face increasing problems of water quality deterioration caused by nonpoint-source pollution. The effects of global climate change are not discussed in this paper but the issues raised and solutions proposed are equally applicable.

In looking at water quality problems which face nearly all countries in the world, we must consider specifically the problems of resource management from the perspective of research, regulation, and action agency needs. All levels of society face problems created by deteriorating water quality. Perhaps the most significant tool available to all these groups is the mathematical model. Obviously the same model will not be satisfactory for each application, and in fact, there are large differences between appropriate models. In this paper we discuss the needs of these managers from a philosophical perspective, identifying what motivates them to use models, the characteristics of existing models, features needed on models now, and ideas for future development and use. Even though users require different levels of development, the philosophical comments apply to all levels. Emphasis will be placed on non point sources of contamination.

[1]Donn G. DeCoursey, Research Leader, USDA-ARS, Hydro-Ecosystems Research Unit, Fort Collins, Co. 80522
[2]James S. Schepers, Soil Scientist, USDA-ARS, Soil/Water Conservation Research Unit, Lincoln, Ne. 68583

MODEL USE

Why use models in the study of nonpoint-source water quality problems? The answers are obvious to most model users, but it serves to introduce some of the issues we want to discuss. Many of the chemical fate processes (like transport, transformation, decay and adsorption) are so interwoven that frequently we cannot isolate and measure them experimentally. It is possible to use mathematical models that simulate the physical processes in a simulation mode to supplement laboratory experiments and gain an understanding or test our ideas of how processes react and interact in real world situations. In some cases, however, mathematical models used to describe different processes can lead to the same analytical expression. van Genuchten et al (1990) discussed the two-site concept of adsorption/desorption (sorption sites governed by equilibrium and sites governed by first order kinetics) and the mobile/immobile concept of water flow to explain diffusion controlled sorption. Both concepts lead to the same mathematical expression, thus it's not possible to use the models by themselves to isolate the specific processes responsible for kinetic adsorption/desorption processes. Very careful experimental laboratory and field research must be combined with the best mathematical descriptions to determine specifically what processes are responsible for the observed phenomena.

Using mathematical models to simulate a complete process forces the model developer to simulate all of the major subsystems or modules that interact to drive the system. Even if the subsystem plays a rather minor role and can be regarded as insignificant, the developer is aware of its existence and must address it to determine the model's sensitivity to that subsystem. Collectively the development of process models helps identify missing pieces, weaknesses, sensitive and insensitive parameters and subprocesses, and set research priorities. Other authors in this symposium have addressed this issue.

Obviously, one reason for model development is to characterize processes where alternative management practices can offset water quality, thus enabling the user to study alternatives and their costs, and find optimal solutions. Parallel with this motivation toward improved water quality, is the need to predict responses of cropping systems to a variety of changing inputs such as climatic sequences, alternative fertilization practices, water and tillage management options or stress situations. Models with features to simulate the physical, chemical and biological subprocesses provide opportunities for the researcher to communicate with farmers and others who may be responsible for potential pollution problems or desire more information about managing their resources.

PROBLEMS OF MODEL DEVELOPMENT

The motivation for development of good simulation models seems so strong that there should be many good simulation models of all water and solute flow processes including subprocesses associated with them. There is an abundance of such models, but there are relatively few really good ones.

Inadequate Data Bases

Models of water and solute movement described for the root zone, groundwater, rivers and surface impoundments are very complex. This is especially true when the detail required to make the models respond to management alternatives is considered. Herein lies part of the problem. Data bases required to validate, parameterize, and drive models have been far outpaced by model development. Costs of data acquisition prohibit the extensive data collection required for good validation (if true validation is even possible). Thus we find models that are poorly tested; and others that have had no validation at all. This is an argument for better coordination between model development and experimentalist researchers and the user, and points to an obvious need

for a more concerted effort in support of integrated model/field studies of specific processes or modules, (see DeCoursey, 1990). It is also an argument for structuring models to take advantage of existing data bases or providing sufficient justification for expansion of generally available data bases.

Poor Planning

Over the years the defense industry has been required to develop large complex models. Thus they discovered, long before the rest of us, the need for well-developed plans before embarking on a model development project. Many of us developed our modeling skills years ago when computers were in their infancy, and since that time we have not had adequate training in software development. Recent graduates are probably better qualified than older scientists, but the comprehensive models being developed now are beyond the capabilities of all but a few of us. Thus it is not surprising that many of the models described in this book are not in use. A model that is going to see continued use must be maintained and must be easily subject to modification and updating.

Models that have this capability require good planning and design to adequately identify and project software life cycle development costs up front. Boehm (1984) describes a constructive cost model that can be used to better project the time and costs of software development. DeCoursey (1988) surveyed many ARS scientists involved in model development, and found a common feeling among those surveyed: the need for better planning and a more structured approach to development. Software such as Excelerator and ProKit Workbench are tools now available to help in this development. See Hebson and DeCoursey (1987) for an application of these concepts in development of a rootzone leaching model.

Poor Software Development Technology

Coincident with a lack of planning are other software development weaknesses that have contributed to problems in many existing models. DeCoursey (1988) and Reynolds et al. (1988) identified common problems which can be summarized as: (1) modelers and users cannot interpret the code; it is monolithic and poorly documented, (2) the models are complex, poorly balanced and are difficult to parameterize, (3) the code is not adequately tested and (4) poor documentation leads to misapplication. Woodfield (1990), in his paper on software development in these proceedings, identifies new technology in software development that should lead to better models in the years ahead.

Models too Complex?

We have alluded several times to the complexity of water quality models;presentations of several other authors substantiate our concern for this problem. Obviously, unscrambling cause and effect in such systems is difficult. As a result many models have a large number of parameters and require cumbersome input data. It is frequently argued that models requiring large amounts of input data and estimates of numerous parameters are nothing more than curve-fitting processes. Others disagree. These arguments are frequently the result of misapplication or interpretation. Predictive models very rarely can be statistically shown to support more than five or six variable and parameter values. This is because the incremental increase in explained variance, as new variables are added, becomes small compared to the unexplained variance. Predictive models, to which this argument applies, have only one major output feature. Thus, those who use this argument in being critical of the number of parameters some models have, have failed to recognize that the water quality models presented here are generally multifaceted; that is, they have several different types of output. This is especially true of research models which may simulate many different processes and show the results as functions of both time and space.

Perhaps more important than justifying the need for certain parameters is the need to recognize, in advance, the objectives for which a model has been or is being developed. This is a requirement of both the developer and user. It is not realistic to assume that a model developed for research can be used at the field level. However, that does not mean that a field level model could not be developed from the research model following sensitivity analyses and adequate consideration of input requirements of the model and user capabilities.

FEATURES OF MODELS THAT CAN RESPOND TO MANAGEMENT QUESTIONS

Emphasis of this paper is on improving models that managers can use to address problems of nonpoint-source contamination of water. For the most part these are models that action agencies can use in dealing with farmers. The source of pollution is generally an agricultural chemical (fertilizer or pesticide applied on rural agricultural land). Obviously, other nonpoint-source contributors to contamination of surface and ground water include golf courses, parks, greenbelts and even domestic lawns; but potential for agricultural contamination is predominant.

Our comments here apply primarily to rectifying management practices responsible for pollution rather than dealing with pollutants after they get into water supplies. Water supply managers working with eutrophic lakes, for example, have different problems, as addressed in these proceedings by Stefen et al. (1990) and Whitehead et al. (1990). Philosophically our comments apply equally well to models such as those described by Stefan and Whitehead, but the recommendations herein for specific model improvements are directed toward the source of the problem rather than the contaminated water.

Many of the models discussed in this book describe in detail the soil physical processes responsible for the fate of nutrients and pesticides. However, only a very few of the models (research or applied) have the ability to consider parameter values as variables that respond to management practices or that are characterized by spatial variability. We recognize that a few models do have some of these features, but most can be improved to some extent. This is the issue we address in the following section of this paper.

Incorporation of Crop Growth Models into Water Quality Models

Evapotranspiration and erosion play dominant roles in determining the fate of nutrients and pesticides. Both of these processes are drastically changed by the extent and kind of surface vegetation. The source of transpired water depends upon rooting depth, size and age of the crop, water content at different depths in the soil profile, and the physiology of the plant itself. Erosion is a function of the degree of protection afforded by the size of the plant and its physical structure (i.e., whether rainfall hitting the plant is conducted down the stem or drips from leaves on the periphery of the plant), as well as a function of the amount of field area protected by the crop. These and many other plant features, some of which respond to or are affected by nutrients and pesticides, need to be incorporated into farm level management models. There are many crop growth models (cotton, corn or maize, wheat, soybeans and others), see Wisiol and Hesketh (1987). Some of these models are physiological, developed to aid in crop or insect research, others are statistically based yield prediction models, and a few have been developed to aid in day to day management of field crops. GOSSYM (Baker et al, 1983), for example, is being run by an expert system, COMAX, COtton MAnagement eXpert (Lemmon, 1986), on some United States farms to aid in cotton farm management. GOSSYM is not specifically designed to aid in water quality research, but does illustrate what can be done in aiding farm management. In general it is much more comprehensive than most models needed for management applications. GOSSYM and other models need to be modified by teams of experts and managers to incorporate into field level models those aspects of crop growth simulators that change with soil, climate, weather, nutrient levels and management options.

Effects of Tillage and Other Process on Soil Physical Features

B.E. Clothier (1990), in an accompanying chapter, describes in detail how infiltration, soil hydraulic properties and soil water movement are affected by plant structure, plant roots, and tillage. Biological processes, tillage and freeze/thaw activity can change the hydraulic features (such as conductivity and infiltration characteristics) of soil by orders of magnitude. Tillage and the reconsolidation effect of subsequent rainfall creates a dynamic soil bulk density. The presence of plant roots also affects soil water content. Since these processes affect retention and infiltration rates so dramatically, they must be incorporated into management models. Many management practices determine the extent of tillage in one way or another, and thus affect hydrologic response. Preferential flow, the movement of water through continuous pores larger than about 0.5 mm, is responsible for movement of nutrients and pesticides deep into the soil profile. The size and extent of preferential flow paths is a function of tillage, reconsolidation, biological activity cropping history, and soil type. Since preferential flow is so important and responds to management-induced changes, it should be included in management level models. Much research is needed in characterizing these processes and in finding the best way to introduce them into the soil's data base.

Couching Response in Probabilistic Terms

Several accompanying papers including those of Woolhiser et al. (1990), Dagan et al. (1990), and Plate and Duckstein (1990), describe the variability in hydrologic response we can expect in any given field. This variability is associated with many physical features (e.g., up to two orders of magnitude variance in soil hydraulic conductivity), and our inability to adequately parameterize simulation models because of lack of perfect knowledge about processes and inputs. Collectively this means that we cannot say with any degree of confidence what the likely response (say leaching depth of a pesticide), is going to be. However, by using Monte Carlo methods and other statistical techniques, we can run the models many times, within the range of likely parameter values, and quantify the uncertainty of the response by estimating its probability distribution. In some cases it would be advisable to compare probabilistic response of different models, because the Monte Carlo application described above defines only parameter uncertainty and presupposes an accurate model. This may not always be the case, thus the desire to compare probabilistic response of different models. Probabilistic types of information are more useful to managers than single values that we can't put much confidence in; managers need the ability to temper their response to fit the situation (see Shaw and Falco, 1990 in these proceedings).

Nutrient and Pesticide Processes

Much research has been conducted on nutrient and pesticide processes and their fate, however, only a portion of this information has been incorporated into management level models. Pesticide manufacturers, for example, are now starting to provide rates for degradation of pesticides by a variety of processes including volatilization, hydrolysis, biodegradation, etc. If we had these features in our models we could better account for the loss. Much of this detail can be transparent to the user and still yield a more sensitive model.

Research and Data Needs

Our comments in the previous discussion of improvements that can be made in application level models might indicate that we have all the answers. That is far from a reality. Even though we have the ability to improve the crop growth, tillage, and chemical features of the models, much additional research is needed. This research needs to be a coordinated effort in field and laboratory experimentation combined with model development to respond as rapidly as possible to weaknesses in simulation systems. For example, how well can we characterize all the major crops? We know that a lot must be done in the area of preferential flow and in soil parameterization.

Data bases need to be developed and geographic information system technology implemented even more than it is now in amassing data needed to drive the models. Expert systems need to be developed to make models easy to use and complexities as invisible as possible to the user.

SOFTWARE IMPROVEMENT

Use of Generic Modules

Where do we begin to improve the software of water quality models so they will better meet the needs of the field level manager? Reynolds et al. (1988) in assessing the needs of plant growth simulation modelers, identified some of the problems and limitations of their models. As discussed previously in this paper; they recommended development of modular generic simulation models. Their recommendation for the use of a generic approach to modeling is based on the fact that the modeler is forced to determine the general properties of the class of systems describing plant growth and to see a plant-specific growth model as a variation on a theme rather than as a separate entity. They quote Jay Forrester, an expert in systems engineering, in support of their argument: "...one should start not by building a model of a particular situation, but by modeling the general class of systems under study. This may seem surprising, but the general model is simpler and more informative than a model of a special case" (Forrester, 1970).

The concepts discussed below, are generally applicable to most modeling problems, including hydrologic models discussed in this book. But with regard to nonpoint water quality models, the concepts are probably more applicable to the structure of plant growth models and soil process simulators that will be imbedded in our water quality models. Given the wide range of management scenarios likely to be considered by model users, a large number of plant growth and soil process simulators will be needed. If these simulators can be obtained by changing parameter sets of a few general classes of models,they would be easier to use. It should also not be necessary for the user to know anything about the parameters; only that the crop and tillage system be identified, the parameter values would then be inserted automatically.

Justification for Use of Generic Modules

In their paper, Reynolds et al., proposed the development of crop or plant growth models along the line of well-engineered and documented generic modules. Their justification of generic module development is:

1) It enables simulation of individual systems belonging to a class of models simply by changing parameter values. This is a major advantage in that model parameterization is easy and could be done automatically with little input from the user. Different models could be handled the same way, but large volumes of code would be required, much of which would not be used most of the time. Thus, the use of generic modules is not only simpler, but, it is also more efficient.
2) It forces the modeler to determine general properties of the class of systems, and thus see individual systems as variations on a theme rather than separate entities. This is a definite advantage with respect to code structure and efficiency.
3) It leads to development of general models that offer simplicity and initially higher information value than special case models.

These comments about plant growth simulators are equally applicable to systems describing the physical structure of soil as it responds to management, thus the concept of a generic approach to soil process simulators is recommended. There are other advantages to the use of a generic structure to both plant and soil process models:

1) Generic modules are easier to verify and validate.
2) Error detection and correction are simplified.
3) It facilitates the interaction of specialists.
4) Modules developed along disciplinary lines are more intelligible.
5) Maintenance is easier.
6) Testing, changing, and interchanging of code are easier.
7) New systems (plant growth and soil process models) can be developed from existing modules.

Figures 1 and 2 are examples showing a conventional diagram of a plant model (in this case tree growth) and the same model with a modular structure respectively. The elements of the model with a modular structure are similar to the conventional system, except that they have been rearranged into physiological modules. It is easy to see why such an arrangement would facilitate the development of simulation models for other crops.

Obviously, with respect to development of management level water quality models, use of generic modules has more application in plant simulation and in simulating the effects of tillage on the physical structure of the soil, than in other processes. This is because of the large number of similar process modules needed. However the concepts can be applied to all model development in general, where numerous, very similar features are to be simulated.

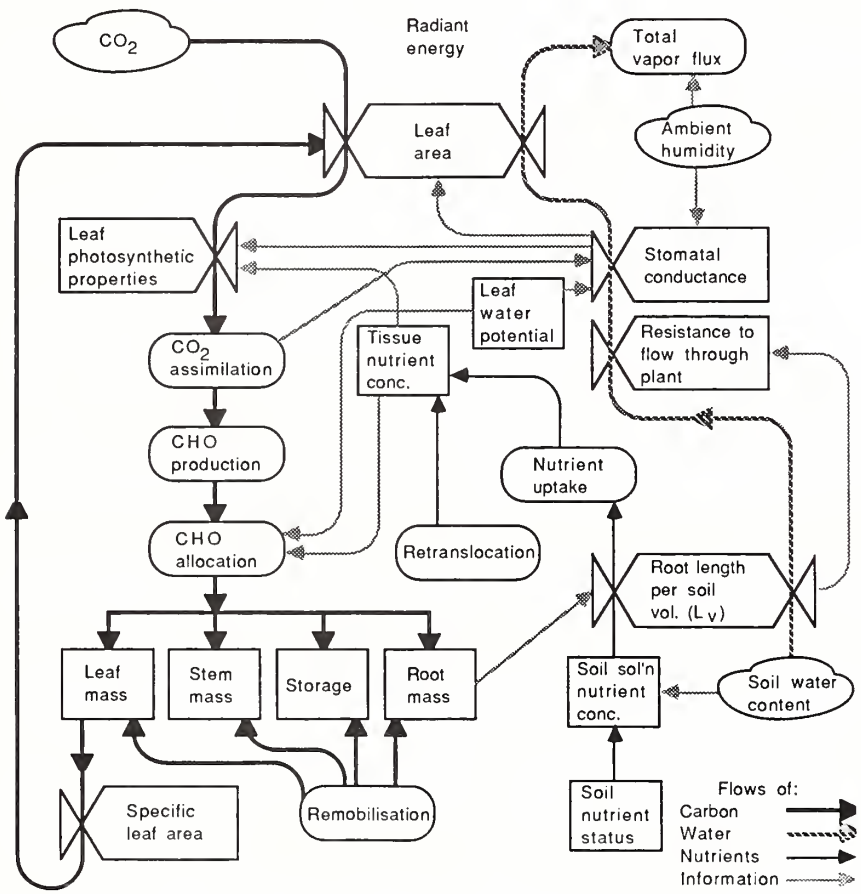


Figure 1.
Diagram of a tree growth model, from Reynolds et. al. (1988).

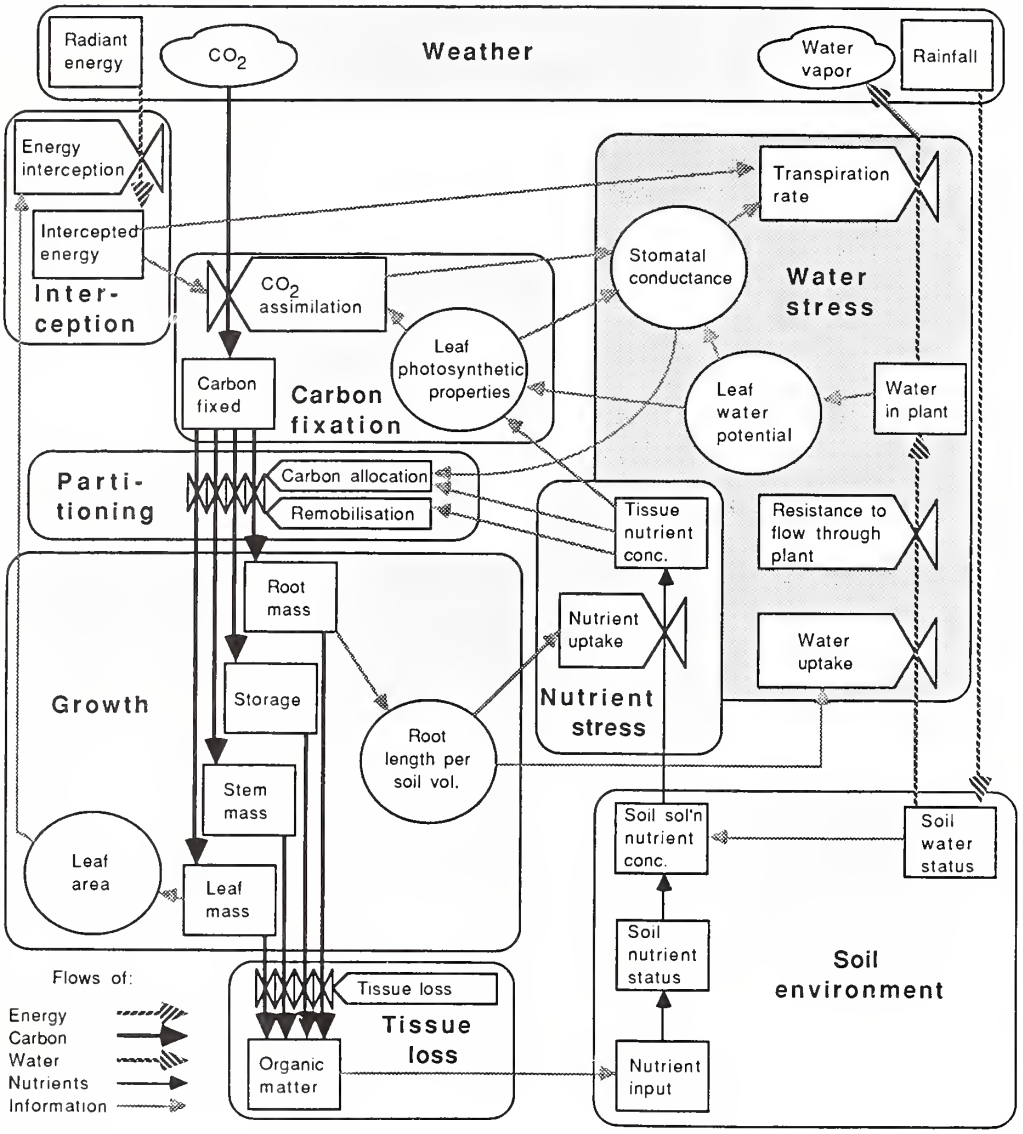


Figure 2.
Diagram of a tree growth model in which the model elements have been rearranged into physiological modules, from Reynolds et. al. (1988).

FUTURE USES OF MANAGEMENT LEVEL MODELS

Our discussion to this point has focused on the types of changes we envision as necessary to improve field level models designed to respond to managers' questions. As additional incentives for development of these models, we suggest that there are other uses that can be made of such models. As these models become more comprehensive, and less empirical, the possibilities for other uses, equally beneficial to the managers, emerge. Examples include:

1) Sophisticated budgeting of water and nutrients
2) Providing guidelines (or restrictions) on fertilizer, water and pesticide applications
3) Synchronizing management practices with climatic patterns and probabilities
4) Defining ideal growth environments in terms of soil water levels, organic matter, and nutrient and pesticide applications
5) Designing and manipulating tillage operations to create an optimal growth media.

These concepts may appear distant, but development is already underway on some of them; applications are being considered in the field of optimal growth environments, for example. Much work at the research level remains before such programs will be generally available. The solutions will not be easy when one considers the vagaries of weather and their impact on production, but in time we will see interaction models that will aid farmers in responding, in an optimal way, to the effects of weather (precipitation, temperature, wind, solar energy, etc.) on productivity, insect infestations, etc.

SUMMARY

In this paper we have attempted to describe some of the problems inherent in field level management models that should be considered by modelers in the immediate future. We describe improvements that could and should be made in modeling physical processes, presenting the output in probabilistic format, and interfacing with data bases. We complete our review with a summary of future uses of such management models.

LITERATURE CITED

Baker, D.N., J.R. Lambert, and J.M. McKinion. 1983. GOSSYM: A Simulation of Cotton Crop Growth and yield, SC. Agric. Exp. Station, Tech Bull., 1908, 4.

Boehm, Barry W. 1984. Software lifecycle factors, pp. 494-518, in Handbook of Software Engineering, edited by C.R. Vick and C.V. Ramamoorthy, 1984, Van Norstrand Reinhold, N.Y., N.Y., 683 p.

Clothier, B.E. 1990. Rootzone processes and water quality: the impact of management, in Proceedings of the International Symposium on Water Quality Modeling of Agricultural Nonpoint-Sources, USDA, ARS-81, In press.

Dagan, G., D. Russo, and E. Bresler. 1990. Effect of spatial variability upon subsurface transport of solutes from non-point sources, in Proceedings of the International Symposium on Water Quality Modeling of Agricultural Nonpoint-Sources, USDA, ARS-81, In press.

DeCoursey, D.G. 1990. Mathematical models: research tools for experimental watersheds in Recent Advances in the Modeling of Hydrologic Systems, NATO Advanced Study Institute, Sintra, Portugal, Kluwer/Reidel Publishers. In press.

DeCoursey, D.G. 1988. USDA-ARS Activities in Hydrologic Modeling, Present and Future, ASAE Summer Meeting, June 28 - July 1, 1987, Baltimore, Maryland, 23 p.

Forrester, V.W. 1970. Systems analysis as a tool for urban planning. IEEE Trans. Sys. Science and Cybernetics 6:258-265.

Hebson, C.S., and D.G. DeCoursey. 1987. Structured analysis/design techniques applied to an agro-ecosystem model. Proceedings 7th Annual Front Range Hydrology Days, H.J. Morel-Seytoux and T.G. Sanders (eds.). Hydrology Days Publications, Colorado State University, Fort Collins, Colorado, pp. 95-109.

Lemmon, H. 1986. Comax: An Expert system for Cotton Crop Management, Science 233, 29.

Plate, E.V., and L. Duckstein. 1990. Stochastic aspects of water quality modeling for non-point sources, in Proceedings of the International Symposium on Water Quality Modeling of Agricultural Nonpoint-Sources, USDA, ARS-81, In press.

Reynolds, J.F., B. Acock, R.L. Dougherty, and J.D. Tenhunen. 1988. A modular structure for plant growth simulation models, in Biomass Production by Fast-Growing Trees, J.S. Pereira and J.J. Landsberg (eds.). Kluwer Academic Publishers, Boston Ma. pp 123-134.

Shaw, R.R. and V.W. Falco. 1990. A management perspective on water quality models for agricultural nonpoint-sources of pollution, in Proceedings of the International Symposium on Water Quality Modeling of Agricultural Nonpoint-Sources, USDA, ARS-81, In press.

Stefan, H., R. Ambrose, and M. Dortch. 1990. Surface water quality models: modeler's perspective, in Proceedings of the International Symposium on Water Quality Modeling of Agricultural Nonpoint-Sources, USDA, ARS-81, In press.

van Genuchten, M. Th., S.N. Gorelick, and W.W.G. Yeh. 1990. Application of parameter estimation techniques to solute transport studies, in Proceedings of the International Symposium on Water Quality Modeling of Agricultural Nonpoint-Sources, USDA, ARS-81, In press.

Whitehead, P.G., L. Somlyody, and G. van Straten. 1990. Surface water quality models for planning design and operational management, in Proceedings of the International Symposium on Water Quality Modeling of Agricultural Nonpoint-Sources, USDA, ARS-81, In press.

Wisiol, Karin, and John D., Hesketh. 1987 Plant Growth Modeling for Resource Management, Vols. I and II, Current Models and Methods, CRC Press, Boca Raton, FL.

Woodfield, S.N. 1990. Object-based software development, in Proceedings of the International Symposium on Water Quality Modeling of Agricultural Nonpoint-Sources, USDA, ARS-81, In press.

Woolhiser, D.A., W.A. Jury, and D.R. Nielson. 1990. Effects of spatial and temporal variability on water quality model development, in Proceedings of the International Symposium on Water Quality Modeling of Agricultural Nonpoint-Sources, USDA, ARS-81, In press.

DISCUSSION OF THE PAPERS PRESENTED IN TECHNICAL SESSION 7, PART 1: MANAGEMENT IMPROVEMENTS IN WATER QUALITY MODELS

Don Jensen[1], Presiding
G. Arthur Shoemaker[2], Recorder

PAPERS DISCUSSED

Root Zone Processes and Water Quality: The Impact of Management by B.E. Clothier

Management Improvements in Water Quality Models by D.G. DeCoursey and J.S. Schepers (substituted speaker for declined paper)

SPECIFIC QUESTIONS AND COMMENTS

Comment: (E. Plate, Institute Fur Hydrologie, Federal Republic of Germany) I am surprised to hear that there is a lack of information on crop growth functions.

Response: (D. DeCoursey, USDA-ARS, Hydro-Ecosystems Research Unit, Fort Collins, Colorado) A lot of work has been conducted on crop growth models; however, most of these have been physiological models for specific crops. More generic models are needed to allow broader application. Specific Application would require defining certain parameters.

Question: (Audience) How do roots influence the hydraulic conductivity of the soil?

Response: (B. Clothier, Plant Physiology DSIR, New Zealand) Roots make up a large portion of the total soil volume. Roots grow and die and they shrink and swell. In situ measurements and rooting characteristics such as root density and root lengths are needed. Much of the information previously gathered on crop rooting patterns has been general. More specifics are needed on the biological aspects of the root development.

Response: (D. DeCoursey) The soil parameters can change as a result of the biological growth and decay of roots and due to the disturbance of the soil by tillage. Within a 2-week period, changes in hydraulic conductivity of up to two orders of magnitude can occur. Field data is needed at the same sites at various times of the year and for various crops to better define model parameters.

Question: (G. Burch, Bureau of Rural Resources, Australia) When gathering field data on soils, is general data adequate versus getting detailed or precise data? On many fields there are considerable systematic variations anyway, so are the detailed specifics necessary?

Response: (B. Clothier) I recommend that we get the detailed specifics in the field to get the basic framework for parameters, then take the data and generalize to fit a particular situation.

[1]Don Jensen, Soil Conservation Service, Salt Lake City, Utah.
[2]G. Arthur Shoemaker, State Conservation Engineer,
Soil Conservation Service, Salt Lake City, Utah.

Question: (D. Jackson, Susquehanna River Basin Commission, Harrisburg, Pennsylvania) Are existing models adequate to predict needed water, fertilizer and pesticide applications?

Response: (D. DeCoursey) Predictions can and are made from existing models; however, the models were not developed for this purpose. The tendency is to over-apply fertilizer. This over-application of fertilizer may add to groundwater pollution. New models are needed to help farmers better define the parameters and consequences of over-application.

Comment: (T. Robertson, USDA-SCS, Washington, D.C.) Will the future water, fertilizer and pesticide models include a surface runoff component in addition to a deep percolation component? SCS needs this type of information when assisting farmers.

Response: (D. DeCoursey) The main emphasis is now on groundwater contamination and crop production. I don't know if runoff is being included in existing studies by Marvin Shaffer. I suggest that you contact M. Shaffer.

Question: (D. DeCoursey) Is it possible to turn the process around and measure moisture deficit in the plant and then correlate that value to a soil moisture deficit?

Response: (B. Clothier) No.

Question: (D. Gustafson, Monsanto, St. Louis, Missouri) Can the crop models consider the variation for various species of a particular crop (such as corn)?

Response: (B. Clothier) Some of the physiological crop models have been developed to handle this type of variation.

Comment: (D. Watts, Department of Agricultural Engineering, Lincoln, Nebraska) In Nebraska, farmers are doing pre-planting soil testing. During the irrigation season, farmers are testing the irrigation water and adjusting fertilizer use accordingly based upon a model evaluation.

Question: (A. Sharpley, USDA-ARS, Water Quality Lab, Durant, Oklahoma) Many farmers use soil tests for determining fertilizer at higher rates. How do we get farmers to apply fertilizer at suggested rates based upon modeling?

Response: (T. Dumper, Soil Conservation Society, USDA, Lincoln, Nebraska) The development of technology is only the first step. We must also develop technical delivery tools and procedures to get this type of information to the farmers.

Comment: (R. Hanks, Soils and Biometeorology, Utah State University, Logan, Utah) The parameter values for models should based upon field measurements. Soil hydraulic conductivity and soil water content (how much water is available for the plant) are two examples of parameters that can be field-evaluated.

Comment: (B. Clothier) I agree with the concept; however, soil water potential is hard to measure in the field.

Comment: (R. Hanks) We must use common sense when gathering some field data. The wilting point, air dry and saturation points can be determined in the field. Other points can be interpreted.

Comment: (B. Clothier) With the use of TDR we are able to get better resolution and can get better values.

A MANAGEMENT PERSPECTIVE ON WATER QUALITY MODELS FOR AGRICULTURAL NONPOINT SOURCES OF POLLUTION

Robert R. Shaw[1] and James W. Falco[2]

ABSTRACT

Applications models are needed to develop national policy and to help field staffs make conservation planning and environmental impact decisions. Natural resource managers are concerned about model complexity and data requirements; output; user friendliness; and completeness in terms of physical, economic, and social factors. Interdisciplinary teamwork by users and researchers is essential for usable models.

INTRODUCTION

Modeling technology is giving us an unprecedented ability to look into agricultural nonpoint-source pollution of ground water and surface water. As managers of resource conservation and environmental protection programs at the federal level we seek further development of applications models for national policy analysis and for regional and local decision making by our field staffs. We need models that cope with the social and economic effects as well as the physical processes of nonpoint-source pollution.

This paper takes the management perspective of the Soil Conservation Service and the Environmental Protection Agency and looks at progress and needs in water quality modeling. It also suggests ways to meet those needs.

OUR MANAGEMENT PERSPECTIVE

Soil Conservation Service

The soil conservation Service (SCS) and the Environmental Protection Agency (EPA) have technical leadership for many of the U.S. Government's natural resource and environmental programs. For both agencies, helping the Nation protect the quality and quantity of its water resources is a high priority.

SCS, an agency of the U.S. Department of Agriculture, is a technical service agency devoted to helping agricultural producers and other land users conserve their soil, water, and related resources. Most of this technical service is provided at the request of the land owner or operator through locally organized and locally run conservation districts, generally at the county level. SCS also makes natural resource inventories and recommends national conservation priorities for the U.S. Department of Agriculture.

Water quality has become a high-priority consideration in all SCS conservation activities. The agency seeks to develop computerized water quality models suitable for onfarm water quality management plans and plans developed for hydrologic units. These models should help assess the movement of pollutants in different soils and site conditions and under different kinds of

[1]Robert R. Shaw, Deputy Chief for Technology, SCS, Washington, DC.

[2]James W. Falco, PhD, Former Director, Environmental Monitoring Systems Lab., EPA; Battelle Pacific Northwest Div., Richland, Washington.

conservation practices. They should also include the economic and social factors needed to determine reasonable and practical conservation alternatives that private landowners will adopt voluntarily. Models also are critical to SCS for national and regional assessments of water quality. These models also must support SCS international technical assistance and they must facilitate coordination between SCS and other agencies like EPA.

Environmental Protection Agency

The Environmental Protection Agency coordinates environmental research, monitoring, standard setting, and enforcement activities at the federal level; and it supports similar state, local, and private-sector activities. For water quality, EPA develops national programs, technical policies, and regulations affecting freshwater and marine environments.

EPA has a major role in determining the environmental impacts of pesticides and in controlling their use by (1) specifying permissible applications, (2) requiring container labels to have factual information and guidance on proper use, and (3) providing guidance to state programs that verify applications of certain pesticides. Water quality models are used in all of these activities. Because the demand for model applications is variable and because specialized expertise is needed periodically for some of these applications, EPA relies on a combination of inhouse staff and contractors to carry out modeling studies. A core of EPA staff must be maintained to respond to demands for rapid technical responses and to ensure that contract studies meet the technical requirements. Equally important, EPA will have to continue to rely on contract support for the periodic specialized expertise needed to cost-effectively manage periods of high demand for modeling studies.

PROGRESS AND NEEDS IN WATER QUALITY MODELING

Managers at SCS and EPA are optimistic that advances in computer modeling of nonpoint-source pollution hold the promise of better policy decisions, better project formulation, and better direct assistance to the public.

Progress

Frequently, modeling is used to stylize plans to help protect water quality. These models span a wide range of complexity, accuracy, and precision. To one degree or another, they take into account such factors as soil characteristics, topography, climate, and structural and management practices in predicting the transport and persistence of agricultural chemicals.

The use of water quality models has increased over the past decade for several reasons: (1) public demand for water resource protection, (2) the need for more precise answers when dealing with many variables, and (3) the improvements in predictive capabilities and greater confidence through field testing and application by larger numbers of scientists and engineers. A number of limitations, however, restrict the use of the available water quality models, particularly in evaluating the environmental impacts of agricultural practices. The manager's needs, in terms of utilitarian models, have not been fully met.

Needs

As managers, we can best discuss our modeling needs in terms of (1) degree of complexity and data requirements of the models, (2) completeness of the models, (3) output, and (4) user friendliness.

Complexity and Data Requirements

Simple and complex models are in use today. Their practical value, in the long run, will hinge on how easily the user can provide the data and the technical expertise needed to operate these models.

Simple Models: The simpler models, such as adaptations of the Universal Soil Loss Equation to predict nonpoint-source emissions from fields, have wide-spread use. While these models are relatively easy to use, they require understanding of the phenomena modeled and a knowledge of the area under study. That means the users must have substantial experience working in the area. They must be able to identify soil characteristics, topographic features, and land use practices that alter parameter values.

Where these simpler models are used, federal managers focus on (1) retaining experts to oversee the application of the models, (2) maintaining an appropriate balance of resources -- from within and outside the agency -- to apply the models in a cost-effective and timely manner, (3) maintaining data files needed to run the models, and (4) collecting appropriate field data at the site of application.

EPA uses the simpler models in a large number of relatively short-term screening evaluations to process the large numbers of requests for pesticide-use permits that come into the regional offices. National policy formulation is another use EPA has for these models.

EPA often hires contractors to apply the water quality models for permit requests. Relatively little effort, however, goes into evaluating the success of these applications. Optimal management of model applications should involve systematic, periodic evaluations of these applications. Such evaluations could define the continuing needs for training, research, and development, and guide how the models are used.

Complex models: By and large, complex models have been used for difficult, longer-term studies and research applications. They could be used, however, in short-duration studies with sufficient advanced planning. That planning involves running many simulations using a range of parameters spanning the expected values. The results of these simulations could be stored in a computer file. For a specific site, the user could search for the simulation most closely matching the characteristics of the site.

In managing complex models, the single most important concern is planning a large enough data collection effort. In many assessments of how pesticides and fertilizers used on fields affect water quality, EPA finds that optimal use of models requires extended high- and low-flow stream records to ensure predictions that are accurate and precise.

From the SCS perspective the data requirement is a primary concern. SCS field offices are rapidly becoming computerized and thus able to adopt reasonably complex models; but, these models should use existing data bases or data that are reasonably easy to obtain (Flach, 1983). SCS field people "cannot take the time to make site-specific studies that require instrumentation or special measurements... therefore, [they] must rely on available information such as standard soil survey data and range site data" (Miller, 1983).

Completeness

As to the completeness of water quality models, our managers look at how well they assess the movement and effects of pollutants, the effects of conservation practices, and the economic and

social decisions that landusers make. How well a model factors in these things, of course, depends on the availability of quality data.

Modeling how pollutants move through the environment is essential for EPA's environmental-impact assessments and for SCS's technical assistance operations. Needed is a model that has separate components for the erosion processes of detachment by rainfall, detachment by runoff, sediment transport by runoff, and deposition by runoff.

Along with models that show the movement of pollutants, we need information on the characteristics of pesticides and on how these chemicals transform in the environment. Such information is essential for the operation of models that attempt to simulate the movement of pesticides through the air-soil-water matrix (SCS, 1987).

EPA foresees a need for sophisticated computer modeling systems as science breaks new ground in water quality assessment. For example, using models to predict the behavior of genetically engineered organisms presents challenges not previously encountered in estimating the transport and transformation of chemicals in the environment. Among those challenges are predicting the growth and death of organisms in water bodies and predicting the possible transfer of altered DNA between individuals of the same species and among different species. To address these challenges, we need data on the growth requirements of altered organisms.

Interpreting temporally varying estimates of environmental pollution is another challenge. At present, toxicological information for two types of exposure situations is commonly reported. The first type usually characterizes acute and chronic effects resulting from exposure to chemicals at high doses over short time intervals. The second type characterizes the effects of exposure to low steady-state concentrations over long time intervals. Intermittent exposures have been evaluated only in a few instances; however, we can expect improved techniques for predicting the effects of such exposures, and we can expect that our managers will have to be trained to use these techniques.

Simulating how pollutants move through the environment, and transform, would help SCS in conservation planning. The physical effects of conservation practices could be anticipated.

SCS needs a water quality model for watersheds of up to at least 400 square miles. This model would predict daily values for "water yield; sediment yield by particle size, including organic matter; and other water quality factors such as water temperature, dissolved oxygen, and chemical movement and transformation" (SCS, 1987).

Improving the way we analyze relationships between land and water on a hydrologic basis should be a goal in technology development. Built into hydrologic models should be the movement of pollutants through soils in a toposequence.

SCS needs models that help assess how much the agricultural practices recommended by the agency affect water quality. Recognizing that different combinations, or "systems", of practices have different effects on groundwater and surface water quality under different conditions, the SCS goal is to understand the environmental and economic effects. Models would help determine these effects by category of impact (such as fish habitat) and by parameter (such as sediment) (SCS, 1987).

As SCS watershed protection projects and other federal programs are required to justify their costs on the basis of offsite effects of agricultural conservation practices, the role of computer modeling increases. SCS associate deputy chief for programs, Dennie Burns explained it at a modeling symposium in 1983:

"We need to do a better job of determining the rate of accrual of benefits and costs associated with both the public and private benefits....We need to do a better job of determining the minimum level of incentive payments needed to demonstrate technology and motivate farmers to assume higher levels of management. We need to put all of these data on costs and benefits together for analysis in project areas and target areas, and even national analysis. We need to be able to determine cost-sharing rates for projects quickly and with less research."

There's practically nothing in the way of modeling that deals with the social impacts of changes in water quality. We need to understand how people are affected, under what conditions they will participate in public programs, and how they react to public policy.

The socio-economic linkage will continue to be relatively untouched until we agree to put a value on certain things. Federal managers must have the modeling tools to deal with many social and economic variables: changes in the size, ownership, and financial status of farms; onfarm technological changes; changes in the social and demographic characteristics of rural communities; and, federal policies. Yes, human behavior is hard to predict; but we can at least try to predict the direction and magnitude of effect. We need to fit variables in models, and we need to make the models usable.

Output

In terms of output, managers in SCS and EPA are concerned about the tradeoffs in computer efficiency and accuracy. We're concerned about errors introduced when we have to lump variability (Robbins, 1983). It would be useful to know the extent of this error and its cumulative effects. It would also help us to know which assumptions generate the most significant error. Hopefully, the time constraints and the costs of data collection, which force us to lump variability, may be overcome as we utilize remote sensing and geographic information systems.

We need to learn more about the probabilistic models that research is starting to produce. These models provide estimates of the probability of the occurrence of environmental concentrations of chemicals over various time periods. Interpreting results from these models will be different from interpreting results from deterministic models. In general, most deterministic models give a single best estimate of exposure levels with consideration given to the uncertainty in this estimate. For probabilistic analyses, no single value estimate is used. As these models come into more common usage, we'll have to train our field people to analyze results.

We want standards to ensure that models are verified and validated. The bottom line is that our models must be robust. "A robust model...is one that is stable under a wide range of conditions. A stable model generates answers that make sense. It does not blow up, that is, does not give nonsense answers...and it does not need extensive manipulation. A robust model is also user friendly" (Miller, 1983).

User Friendliness

More could be done to make our research models user friendly -- useful to practitioners. We need to package the models currently available for use in the field.

Those who develop the models must understand that user agencies have varying demands; therefore, a single-applications model is not practical. Research must produce a research model that is robust enough for a broad range of managers and staff; this requires testing in the field. This research model then forms the root for various applied models that cover a wide range of user needs.

SCS field staff will not be running models of the sophistication to which most modelers are accustomed. They need efficient ways to run a series of landscape conditions that describe a variety of choices for resource management systems.

IMPROVING TECHNOLOGY TRANSFER IN WATER QUALITY MODELING

Technology transfer in water quality modeling requires a coordinated, interdisciplinary effort. Interdisciplinary teams of model makers and users must coordinate the following:

- Defining end-user requirements. SCS, for example, has to provide landusers with conservation alternatives; SCS models, therefore, have to be user friendly and less sophisticated than the researcher's.

- Ensuring interdisciplinary and interagency communication. In other words, we need to all "speak the same language."

- Distinguishing one model from one another. Is one model better than another? Which is best theoretically? Which is best from a practical standpoint? Which are the best parts of each? Can the best parts be combined? If not, why not? If we're proliferating the same models in different programs, why so?

- Meeting data requirements:

  "For models developed in the future, we need enough lead time to begin assembling the required data. With the extensive data required to run most of the research models, we must be concerned with the sensitivity of the model results to variations in the input data. If we need data we are not collecting, we must collect them. If we are collecting data we do not need, we should stop collecting them. We should also spend the most effort collecting the data that are the most important in terms of model sensitivity" (Okay, 1983).

- Ensuring that operational models are "...validated and the computational procedures verified so we can determine the sacrifices in accuracy for gains in efficiency. A comprehensive validation procedure is needed at every step because a simple model built on experience and judgment but not fully validated is not ready for use even though it could be easily implemented at the field level" (Okay, 1983).

- Making the best use of geographic information systems and expert systems. Improvements are underway that can make complex computer processing functions easier to use by applications-oriented specialists. Perhaps the most significant improvements are geographic information systems and so-called "expert" or "artificial intelligence" systems.

  Geographic information system (GIS) technologies show great promise for pulling together and analyzing large volumes of data from a variety of sources -- and without the problems associated with averaged or "lumped" values. What SCS needs are new computer-based models fully integrated with GIS capabilities to evaluate on/off-site effects of soil, water, and wind erosion, water quality, sedimentation, and contamination under alternative management strategies" (SCS, 1987).

  SCS is now testing GIS software called GRASS (Geographical Resources Analysis Support System) and its compatibility with SCS's field office CAMPS (Computer Aided Management Planning System) software. What SCS sees as the bottom line is: GISs will handle large volumes of spatial and nonspatial data, improve the design of models by making geographic

data and analytical tools available, provide map output, and facilitate the use of expert systems and decision-support software for better user friendliness.

Expert systems should be able to help us build in the functional relationships needed in water quality models. Because the staff using these systems will have limited experience, they will need guidelines that clearly define the circumstances under which an expert system can be used and explain the limitations of results generated through the use of the system. This kind of guidance should minimize inappropriate use of the models and misinterpretation of results, two common occurrences in application of water quality models.

- Providing training in the use of models for field operations. Clearly, training will be an important aspect as expert systems are brought into use. Staff who use expert systems will be exposed to a different approach to resolving environmental issues. If managers expect their staff to use expert systems in appropriate situations, training will have to produce not only competence in using these systems but also managerial confidence in the results from expert systems analysis.

SUMMARY

The Soil Conservation Service and the Environmental Protection Agency need water quality models for conservation planning and for assessing the impact of agricultural chemicals. We are concerned with the complexity and enormous data requirements of some of the available models; their completeness in terms of physical, economic, and social factors; their reliability of output; and how user friendly they are.

The computer-based models of tomorrow will most likely incorporate artificial intelligence techniques. Likely also, they will be fully integrated with geographic information systems to evaluate the spatial effects of alternative management strategies on water quality.

Cooperation between user agencies and researchers is essential if we expect to develop models that help us make practical, site-specific decisions from large and diversified sets of data. Care in developing applications models, and in verifying and validating them, should pay dividends in better environmental protection and more efficient use of the public conservation dollar.

LITERATURE CITED

Burns, D.G. 1983. "Erosion and Sedimentation -- SCS Overview," pp.421-422, Proceedings of the Natural Resources Modeling Symposium, Pingree Park, Colorado, October 16-21, 1983, U.S. Dept. Agric., Agric. Research Serv., ARS-30, April 1985.

Flach, K.W. 1983. "Erosion Processes--SCS Issues and Concerns, p.424, Proceedings of the Natural Resources Modeling Symposium, Pingree Park, Colorado, October 16-21, 1983, U.S. Dept. Agric., Agric. Research Serv., ARS-30, April 1985.

Miller, N. 1983. "Overview of SCS Requirements," pp.329-330, Proceedings of the Natural Resources Modeling Symposium, Pingree Park, Colorado, October 16-21, 1983, U.S. Dept. Agric., Agric. Research Serv., ARS-30, April 1985.

Okay, J.L. 1983. "ARS/SCS Natural Resources Modeling Symposium -- Summary Comments of SCS," p.130, Proceedings of the Natural Resources Modeling Symposium, Pingree Park, Colorado, October 16-21, 1983, U.S. Dept. Agric., Agric. Research Serv., ARS-30, April 1985.

Robbins, S.W. 1983. "SCS Overview on Spatial and Temporal Variability," pp.287-288, Proceedings of the Natural Resources Modeling Symposium, Pingree Park, Colorado, October 16-21, 1983, U.S. Dept. Agric., Agric. Research Serv., ARS-30, April 1985.

Soil Conservation Service, U.S. Department of Agriculture, 1987, 1987 Soil and Water Conservation Research and Education Progress and Needs.

DISCUSSION OF PAPERS PRESENTED IN TECHNICAL SESSION 7, PART 2: THE MANAGER'S PERSPECTIVE

Don Jensen[1], Presiding
G. Arthur Shoemaker[2], Recorder

PAPERS DISCUSSED

A Management Perspective on Water Quality Models for Agricultural Nonpoint-Sources of Pollution by R.R. Shaw and J.W. Falco

Water Quality Modeling: The Manager's Perspective by D.G. Jamieson, P.E. O'Connell and G. de Marsily (no paper available)

SPECIFIC QUESTIONS AND COMMENTS

Question: (F. Rittig, Crop Protection Division Product Safety, Federal Republic of Germany) Should monies be spent for single application models versus spending money for the development of general purpose models? (Example: model for determining a power plant location).

Response: (D. Jamieson, Thames Water Authority, United Kingdom) In special cases, site specific modeling can be justified. If a model is needed to make the right decision, the money spent on the model is worth it.

Response: (R. Shaw, USDA-SCS, Washington, D.C.) The SCS is generally more interested in general models. The SCS is generally using models to do more broad based resource planning.

Response: (J. Falco, US EPA, Washington, D.C.) The new chemical models are generalized and are used over and over; however, EPA has done some single application modeling.

Question: (K. Seip, Center for Industrial Research, Norway) Who is the decision-maker on water quality issues?

Response: (R. Shaw) SCS provides technical data to farmers and ranchers and explains the causes and effects of various choices or alternatives. The landowner is the decision-maker.

Response: (D. Jamieson) In my case, the Administrator of State is the decision-maker.

Response: (J. Falco) In EPA, a politically appointed official is the decision-maker, subject to public reviews.

Comment: (G. Burch, Division of Land and Water Research, Australia) In regard to GIS, one must make a distinction between data and information. There will be pressure on analysts to use data that is now being collected.

[1]Don Jensen, Soil Conservation Service, Salt Lake City, Utah.
[2]G.Arthur Shoemaker, State Conservation Engineer,
Soil Conservation Service, Salt Lake City, Utah.

Comment: (R. Shaw) It will take several years to get GIS capabilities incorporated throughout SCS at the county level. GIS is hungry for data. Where possible, we need to use existing data rather than the wholesale gathering of new data. The existing data needs to be reviewed and the holes or gaps in data filled as needed.

Question: (G. Burch) What kind of effort or actions are being undertaken to use GIS-type information?

Response: (R. Shaw) The SCS is just getting started into using GIS. The SCS is looking at GRASS for field office application.

Response: (D. Jamieson) They are spending 150 million dollars over the next 5 years to get into GIS.

Question: (M. Lee, Illinois Sate Water Survey, Champaign, Illinois) Do we need a GIS base model?

Response: (D. Jamieson) Models are now being developed that include GIS as a component.

Response: (R. Shaw) New models developed by SCS will include GIS components. This will force agencies to work together to obtain GIS coordination.

Comment: (D. Jackson, Susquehanna River Basin Commission, Harrisburg, Pennsylvania) I recommend that two additional purposes be added to Robert Shaw's paper:

1) Modeling can be used by River Basin Commissions for a state and regional priority setting and funding of projects.

2) Modeling can be used to determine the impacts the installation of resource management systems have on the individual farm waters as well as downstream receiving waters.

All people do not need to be capable of running complex computer programs. Charts and graphs can be computer developed and made available to people for field use and application.

Comment: (B. Roaza, Florida Department of Environmental Regulation, Tallahassee, Florida) I recommend a symposium on modeling, GIS and Expert Systems be held for managers.

Response: (D. Jamieson) There is a technical lag between new graduates and older managers. In the long run, more managers will have computer knowledge.

Question: (D. DeCoursey, USDA-ARS, Hydro-Ecosystems Research Unit, Fort Collins, Colorado) What do managers see as the role of the universities?

Response: (D. Jamieson) I work directly with university people. I am trying to put together a program to fuse the academic with on-the-job experience.

Response: (R. Shaw) The universities have a role in the development of models and assisting agencies to adapt models for specific needs.

Response: (J. Falco) The EPA has used several universities to develop various models. Technology transfer is a major role of the universities.

THE POTENTIAL IMPACT OF SOFTWARE ENGINEERING ON WATER-QUALITY MODELING

D.G. Jamieson[1] and K. Fedra[2]

ABSTRACT

Low-cost, interactive computing with advanced color graphics is already transforming the way in which water-quality problems can be addressed. Recent developments in software engineering such as Geographic Information Systems and Artificial Intelligence should further enhance this capability, culminating in decision-support systems in which the model is largely transparent to the user.

INTRODUCTION

Even in those countries where much of the gross pollution of water-courses has been removed, there frequently remains a problem with agricultural nonpoint sources such as nitrates, phosphates and certain trace-organics which cannot be eliminated by further capital investment. During recent years, this problem has been exacerbated by the higher river quality standards that are necessary to meet increasingly stringent legislation on potable water supply. By their very nature, these nonpoint sources are diffuse, capricious and ill-defined. In many instances, they cannot be measured directly and have to be inferred from their observed effect, often by a process of elimination. Such are the problems of even estimating the amounts entering the aquatic environment, it is perhaps not surprising that most water-quality studies to date have concentrated on point-source inputs.

However, the need is for modeling techniques which are capable of including both point and nonpoint sources in a realistic way. Their role would be to predict the consequences, evaluate various options, identify the preferred solution and establish its robustness to uncertainty. These models would have to be time-variant in order to cope with seasonality and longer-term effects. Moreover, the spatial variability of the inputs is likely to dictate a distributed or semi-distributed representation. Bearing in mind the stochastic nature of the problem, these requirements inevitably point toward the use of simulation.

Simulation tends to be a cumbersome procedure which requires copious amounts of input data. In the case of agricultural nonpoint sources, data acquisition is normally a tedious and time-consuming activity since historic records such as fertilizer application rates invariably reside on external databases if they have been computerized at all. Furthermore, simulation is an evaluation technique rather than a planning process and therefore has to be coupled with some form of decision-making capability, subjective or otherwise. These and other similar considerations suggest that there is ample scope for improving the ease of application and the degree of objectivity.

---

[1]D.G. Jamieson, HQ Operations Manager,
Thames Water Authority, Reading, England.

[2]K. Fedra, Project Leader, Advanced Computer Applications, International
Institute for Applied Systems Analysis, Laxenburg, Austria.

Fortunately, progress on the computing side has not been restricted to hardware development, although that in itself has had a profound effect on water-quality modeling over recent years. With the cost of hardware falling, manufacturers are now having to concentrate more of their resources on software, not only to protect their market share but also to augment their income. These developments offer prospects of reducing the amount of effort and the degree of skill needed to apply the latest modeling techniques, particularly in terms of model formulation, data manipulation and process optimization.

TRADITIONAL APPROACH

The traditional approach to large-scale modeling of agricultural nonpoint sources is perhaps adequately represented by the Thames Basin nitrate model (Sinnott and Jamieson 1982). In this study, a preliminary assessment based on time-series analysis had indicated that there was a real threat to the Authority's ability to comply with the European Economic Community's Directive on the maximum admissible level of nitrates in potable supply (Onstad and Blake 1980). The next stage was to develop a means of evaluating alternative strategies for meeting the imposed standard. While it was possible to reduce the amount of nitrates in potable supply by denitrifying abstracted water, improving upstream effluent quality, increasing the amount of reservoir storage, and improving the mixing characteristics of reservoirs, the effectiveness (and for that matter cost) of these actions varied significantly in different circumstances. What seemed to be lacking was an objective way of deciding the appropriate mix of alternatives, since it was unlikely that any one option would be superior in all circumstances.

To that end, a planning methodology was developed which took the form of stochastic simulation. In essence, this was an extension of the approach adopted for the quantity aspects of water-resources planning (Sexton et al. 1979). The water-resources quantity model was used to generate the hydrological inputs to the nitrate model. To this was added the nitrate component resulting mainly from agricultural practices and to a lesser extent, sewage effluent. The model purports to imitate reality inasmuch as the concentrations undergo changes and delays as nitrates pass through the system.

A modular structure was used to formulate the overall model in which component models representing the various processes involved could be linked in accordance with the physical system. Five basic component models were identified namely:

(1) soils subsystem (providing for different soils, crop types and agricultural practices);

(2) aquifer subsystem (simulating the movement of nitrates through porous media);

(3) channels subsystem (allowing for adsorption, longitudinal dispersion etc.);

(4) reservoirs subsystem (representing the effects of storage); and

(5) treatment subsystem (comprising a variety of processes for reducing nitrates).

The regional water resource system, as modeled to accommodate water-quality parameters, is very large and extremely complex. Besides allowing for the interactions between quantity and quality, the model differed from that for quantity alone in terms of spatial resolution. Since nitrate concentrations change with time and space, the delineation of the main river channel required significantly more detail than the simple delay and attenuation functions adopted for flow routing. Similarly, although many of the reservoirs are interconnected and could be modeled as a lumped

system from a water quantity standpoint, water quality considerations necessitated a more distributed representation to preserve the management options such as blending water of different nitrate concentrations.

Having adopted a modular structure to modeling, it was a relatively simple task to insert new resource options, alternative treatment processes, etc., anywhere within the existing system and to assess their performance in terms of maintaining the level of nitrates in potable supply below the EEC standard, subject to a tolerable frequency of failure. Thereafter, for all feasible solutions identified which met the required performance criteria, it then became a matter of analyzing the cost of each, with a view to determining the overall minimum cost, both capital and operating.

It will be no surprise that a study of this magnitude consumed a considerable amount of time, not to mention mainframe computing resources. Detailing the actual system configuration on the ground had previously taken almost six months to complete. The subsequent model formulation took twelve months. Data acquisition was scheduled for the same period as model formulation since there was a degree of interaction when it became apparent what records existed. The spatial resolution of the model was largely a matter of previous experience. Similarly, the search for feasible solutions was based on trial-and-error rather than optimization and this again depended heavily on user expertise. Moreover, the costing of feasible solutions was treated as a separate exercise instead of an integral part of the evaluation procedure. In conclusion, while the overall study was judged to be a success, no one was anxious to repeat the experience.

RECENT DEVELOPMENTS

In recent years, there has been enormous progress in making substantial computing powers available to users at relatively modest costs. This increasing availability of computers, coupled with recent advances in software have transformed the way in which water-quality problems can be addressed. These changes in computing technology have already been extensively documented elsewhere (Fedra and Loucks 1985, Fedra 1986, Loucks et al. 1985). It suffices here to highlight the more important developments which will affect water-quality modeling directly. These include interactive computing, computer graphics, relational databases, geographic information systems and artificial intelligence, which collectively have been referred to as software engineering, for the purposes of this paper.

It is sometimes hard to believe that interactive computing has only come to prominence during the 1980s, largely as a result of the phenomenal growth in personal computers. Prior to that, batch processing on mainframe computers was the accepted norm for most modeling applications. The benefits of interactive computing whether it be on a mainframe, specialized work-station or micro-computer are significant in terms of program development and testing. However, it is the so-called 'user-friendly' software available on personal computers which is having the most influence on water-quality modeling through the provision of a more helpful, responsive and flexible computing environment. This in turn is leading to a change in the type of person using modeling techniques. No longer will it be necessary for those concerned with decision-making to rely on a third-party specialist to formulate a model on their behalf and identify the range of feasible solutions. It is becoming increasingly likely that they themselves will attempt to structure their own problem, prompted by expert guidance and easily understood instructions.

Much of the success associated with personal computing can be attributed to color graphics. Bit-mapped graphics with multiple-window capabilities allow the structuring of complex displays. This can greatly increase the amount of information communicated, and at the same time improve the ease of understanding (Meyrowitz and Moser 1981). Mixtures of alphanumeric, symbolic and graphical elements using familiar backdrops such as maps or flowchart representations, can be very

effective. They do, however, require a considerable amount of design effort (Loucks et al. 1985) which needs to be considered from the onset. Nevertheless, if these facilities are built into proprietary software packages for modeling water quality, such considerations are only of passing interest to users.

Since most water resources data are spatially located, Geographic Information Systems (GIS) are likely to gain increased prominence during the coming years. The concept comprises a database of assets and asset-related data which are underpinned by a digital mapping system. Access is obtained by specifying which entities are to be retrieved, together with the criteria for selection which can either be geographic (e.g. within a given area) or attribute based (e.g. river reaches with the same water-quality objective) or a combination of both. Pictorial representation of the geographic data can be scaled and entities selected by means of pointing a cursor. Related data which may be alphanumeric or image, can also be displayed on the screen. When used in association with a relational database management system, these techniques provide an extremely powerful set of tools for manipulating data prior to modeling, for example as in figure 1 (Annand 1988).

Application programs such as water-quality models can be interfaced directly with GIS, thereby providing an efficient means of marshalling the necessary input data. However, that belies the problem of formulating the model in the first instance. At the present time, model formulation is subjective and heavily dependent upon previous experience. Clearly, there is a need for specialist advice, not only in model formulation but also interpretation of results, and to that end, Artificial Intelligence in general or Expert Systems in particular may find a role.

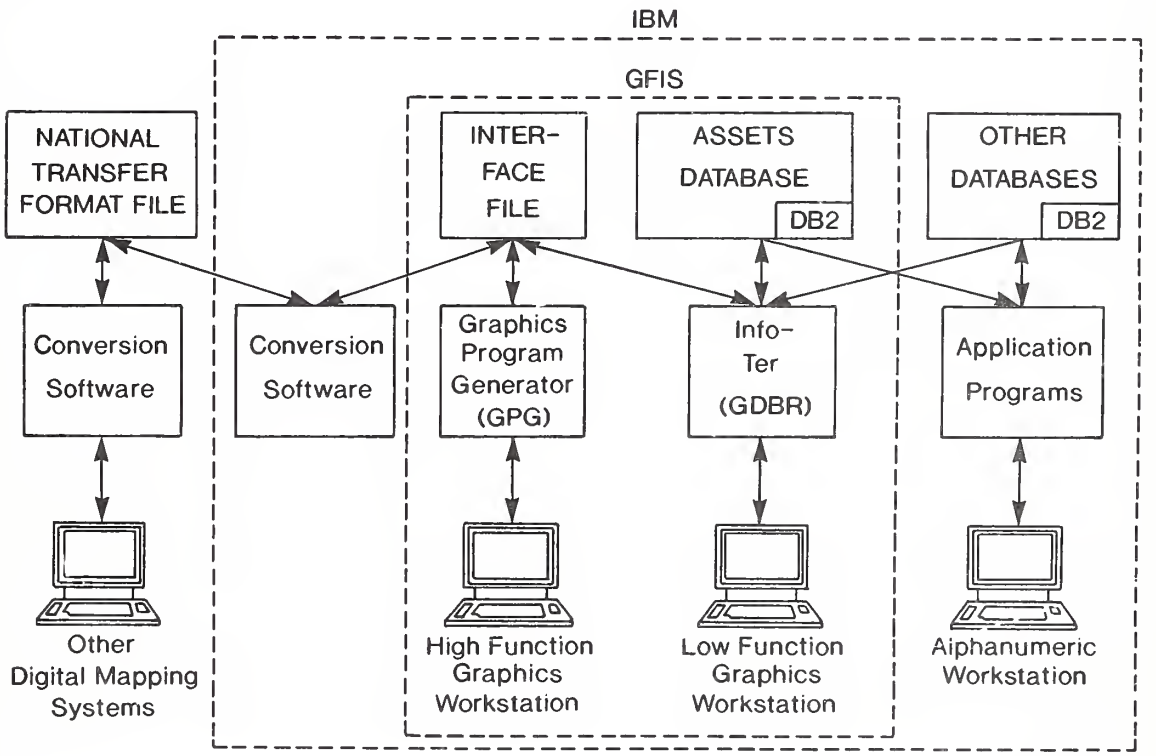


Figure 1.
Use of a geographic information system for asset management.

Expert Systems (ES) offer the prospect of improved support for problem-solving in areas which hitherto have not been amenable to conventional analytical techniques. These problems are typically ones requiring the application of domain knowledge rather than numerical algorithms. The distinctive features of ES are firstly, the ability to draw inferences from a set of supplied information and secondly, the capability of explaining those inferences, solutions and recommendations to the user. In general, ES can be distinguished from normal data-processing (DP) systems as follows (Slinn 1988):

	ES	DP
(i)	Primary symbolic	Primarily numeric
(ii)	Knowledge-based	Algorithms/data-driven
(iii)	Problem structured	Program structured
(iv)	Heuristic (implicit solution steps)	Deterministic (explicit solution steps)
(v)	Rules and facts	Files, records and variables
(vi)	Declarative programming	Procedural programming
(vii)	Control structure separate from domain knowledge	Control structure and data interspersed

Figure 2 shows the ES structure which essentially comprises three sub-systems, namely user interface, inference engine and knowledge base. The user interface allows the system developer to input rules/knowledge, set up search mechanisms/control strategies and provide menu-driven user-access. The inference engine which is commonly supplied by a shell program, takes the user input and provides guidance/solutions/recommendations by inference or interpretation of the rules/knowledge in the knowledge base. The latter is a data store containing domain knowledge which can be either rule-based or frame-based.

PROSPECTS

Given the recent advances in software engineering, it is of interest to speculate how the Thames Basin nitrate model might be approached if it were necessary to repeat the study at some time in the future. For the purposes of this exercise, it is assumed that a technical work-station is available, featuring a high-resolution color-graphics screen, several megabytes of directly-addressable RAM, several hundred megabytes of high-performance mini-disk drives and coupled to a host mainframe computer via a wide-area communications network.

The application program comprising proprietary software packaged as an Expert System for interfacing with the user, would reside on the work-station. This could draw on the host computer for corporate information by means of the Geographic Information System and copy onto its own mini-disks. Definition of the physical system would present no difficulty since that is contained in the asset database. All that should be necessary is to define the scope and degree of resolution, the latter being advised by the ES after consideration of the ostensible objectives and the data availability.

Model formulation would be conducted interactively. Communication with the user should be screen-based, making extensive use of symbolics and color graphics. Sub-system models representing the various components such as rivers, aquifers and reservoirs would be configured by the user in an order corresponding to the physical system. Throughout this period, the user would be guided by the explanation/justification capability of the ES.

Having defined the model structure, the necessary input data would be assembled on the host computer, courtesy of GIS. These comprise various types of asset-related data including

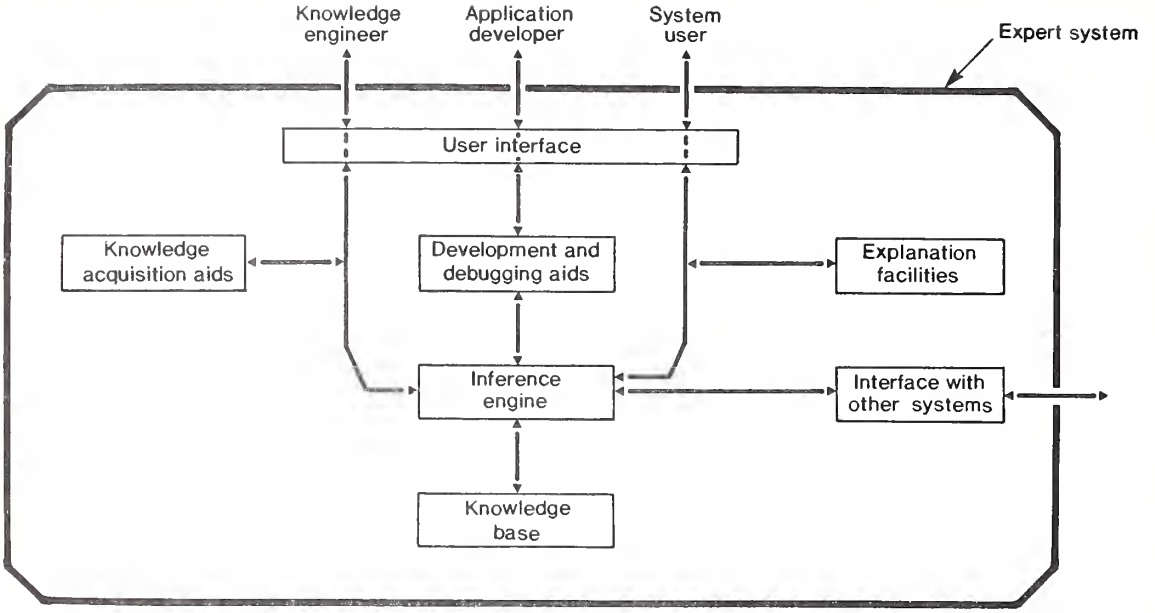


Figure 2.
Expert Systems Structure.

capital/operating costs, measured physical parameters, and time-series, both historic and projected. External data such as fertilizer application rates are likely to be entered onto the mainframe database via an applications program. Similarly, any preprocessing required to interpolate for missing data, correction of boundaries etc. could be undertaken by a facilitating program supporting GIS. Either way, the specified data needs, correctly formulated and indexed, would be made available to the water-quality model with minimal fuss and intervention.

Thereafter, fitting the model by means of parameter estimation would be automatically undertaken and verified by a split-record test. In place of the trial-and-error approach, the search for the optimal development strategy would be conducted in a structured way. Identification of feasible solutions and cost optimization would be performed concurrently rather than as two separate stages. Previous attempts of achieving this end have indicated that while an efficient numeric optimization package can be developed for a specific model, its portability was limited (Smith 1977). This and other considerations suggest the search procedure would be based on heuristic optimization within the ES.

Finally, after the preferred solution has been tested for robustness against uncertainty, the ES can assist the user with interpreting the results from a practical standpoint. In this way, the ES not only acts as a user-friendly interface but also supplements user experience in an important but often neglected area.

CONCLUSION

With the advent of cheaper computing, more power can be afforded in the creation of a man-machine interface which is tolerant of the human idiosyncrasies associated with real-world decision-making. While in the past, interaction with computers was designed to make matters easy for the machine, now the reverse is taking hold. The new software systems which are becoming available no longer impose a formalized structure. Moreover, increasing emphasis is being placed

on ease of application. As a result, these new systems are likely to be less rigid, less time-consuming and less demanding in terms of user experience.

The comparative ease of use will extend the attractiveness of modeling techniques to a wider audience. Whereas previously, modeling was regarded as a specialist activity where more often than not, the person involved had created the model in the first instance, soon there will be a whole class of users who are simply applying established techniques. This has its dangers inasmuch that inexperienced users will not necessarily appreciate the limitations of the model. After all, models are only caricatures of reality and therefore it follows that at best, their outputs are approximations which should not be treated with undue reverence.

Given time, advances in software engineering will reduce but not eliminate this problem. Within the foreseeable future, computers will be used to supplement user experience. In this type of environment, it seems likely that the importance of modeling techniques will be relegated to that of an analytical kernel buried in a decision-support system, with emphasis placed on the quality of the decision rather than the elegance of the model. Eventually, the actual mechanics of deriving the preferred solution could largely become transparent, providing the user understood the system logic. Either way, the initial concern of the user would be to describe the problem and state the objective. Thereafter, the decision-support system should provide all the assistance needed.

REFERENCES

Annand, K.P. 1988. A geographic information system for river management. Automated Mapping/Facilities Management Conference, Nottingham, England.

Fedra, K. 1986. Decision making in water resources planning: models and computer graphics. International UNESCO Symposium on Decision making in Water Resource Planning, Oslo, Norway.

Fedra, K., and D.P. Loucks. 1985. Interactive computer technology for planning and policy modeling. Water Resour. Res. 21:114-122.

Loucks, D.P., J. Kindler, and K. Fedra. 1985. Interactive water resources modeling and model use: An overview. Water Resour. Res. 21:95-104.

Meyrowitz, N., and M. Moser. 1981. BRUWIN: an adaptable design strategy for window manager/virtual terminal systems. Proceedings of 8th Annual Symposium on Operating Systems Principles (SIGOPS), Pacific Grove, California.

Onstad, C.A., and J. Blake. 1980. Thames Basin Nitrate and Agricultural Relations. Proceedings of the ASCE Symposium on Watershed Management, Boise, Idaho.

Sexton, J.R., D.J. Cook, and A.E. Jones. 1979. Water Resources model: An introduction. Thames Water Authority publication, London England. 10 pp.

Sinnott, C.S., and D.G. Jamieson. 1982. River basin planning for control of nitrate pollution. Water Sci. Tech. 14:245-252. Capetown, South Africa.

Slinn, M.A. 1988. A short-term Plan for the Introduction of Expert System Techniques in Thames Water, Internal Report.

Smith D.K. 1977. An approach to the optimization of simulation models of water resource systems. Ph.D. thesis, Dept. of Operational Research, Lancaster University, England.

OBJECT-BASED SOFTWARE DEVELOPMENT

Scott N. Woodfield[1]

ABSTRACT

Over fifteen years ago the basic principles of object-based or object-oriented software developed were described. Now these principles and concepts are supplanting the older function-based methodologies. It appears that object-oriented design will be the methodology of choice in the future. The use of these newer concepts significantly lowers maintenance costs, reduces development time, makes it easier to understand software, and increases the quality of the software. This paper presents the ideas behind object-based software development, describes an object-based design methodology, and presents definitions and guidelines for developing good designs.

INTRODUCTION

The use of higher level languages and good software engineering practices has produced better software than that created in the 1960's. Unfortunately, many problems still exist. It is still difficult to create and maintain quality software. Software is still late and over budget and in most cases does not meet the user's needs. In addition to these perpetual problems, owners of existing software must modify their code very cautiously because of the possibility of doing irreparable damage.

The crux of these problems is failure to adequately model the real world and failure to build software that can adapt to changes in reality. Instead of building software that mirrors reality, most developers shoehorn reality into an awkward, inflexible, difficult to understand, programming paradigm. Specifications and design should reflect reality rather than making reality adjust to someone's programming model.

When reality is described verbally we talk about entities, actions, and events. Most current languages and development methodologies force us to think primarily in terms of actions or tasks and to ignore entities and events. For instance, one of the most popular methodologies, structured analysis and design (Myers 1978, Yourdon and Constantine 1979) is based primarily upon the idea of functions. For instance, in data flow diagrams the primary construct is the transform bubble. In the structure chart the primary construct is the module. It is true that in both cases data is represented but it is essentially a secondary concept. Notice that events are not explicitly represented in the standard "structured" methodologies. Only the simulation community considers events as a first class concept. Unfortunately, entities appear to be conceptually unimportant. While programs contain many variables, there are only a few different variable types. Everything is classified as an integer, real, character, Boolean, array, record, set, or file. Programs written under this constraint are similar to English papers written using any number of verbs and pronouns but only using four or five nouns to represent all conceptual entities. Unfortunately, many computer languages encourage us to write software in a similar manner. We can create any procedure or function desired and any number of variables but are given access to only a few actual types. Even the type definition mechanisms of modern languages do not permit the creation of new "true" types. A true type definition defines not only the set of values associated with a variable but also the set of valid operations. For instance, we understand the type name "character" to not only mean the set of characters that can be manipulated but also the set of operations that can be performed on characters.

[1]Scott N. Woodfield, Associate Professor, Brigham Young University, Provo, UT.

We know the "+" operation is not valid for the type "character". Standard type definitions only define new sets of values. The set of operations that can further clarify our understanding of a new type cannot be explicitly associated with the new type. A new concept, object-classes, has been invented to allow the definition of types in terms of value sets and operation sets.

Modeling reality poorly is not the only cause of current software development and maintenance problems. Lack of information localization is another cause. When creating or understanding a piece of software all the pertinent information about an abstraction should be easily accessible. Functions and procedures make all the information about an operation accessible by storing it all in one place, the procedure or function definition. Unfortunately, many languages do not support the localization of all information about an entity. Modification becomes difficult when information is not localized. For instance, assume a list is represented by an array and pointers into the array. Also, assume all operations on the list (throughout the program) directly manipulate the array and pointer representation. If the representation of the list is changed from an array to a linked list then every operation in the program that uses the list needs to be modified. There is a high probability that at least one list-access operation would be missed or incorrectly modified. Localizing all of the information about the implementation of a list in one place would minimize many of the modification problems.

A third problem, closely related to lack of localization, is lack of information hiding. When all information about a process or entity is localized only the minimal amount of information needed to use the process or entity should be made available. This minimal set of information is often called the interface. All other information should be made inaccessible and preferably hidden. Most high level languages provide this capability for operational abstractions. The interface to an operation is given using a procedure or function header. The body represents the information to be hidden. Most common languages do not provide such a mechanism for abstract entities.

Localizing the information about entities and hiding the implementation details allows the development of powerful entity abstraction mechanisms that are just as important and useful as operational abstractions. Software development based on this concept is called object-based software development (Booch 1986, Wegner 1987).

OBJECT-BASED SOFTWARE DEVELOPMENT

Objects and Object-Classes

An "object" is defined to be an entity that contains both value information and information about the entity's behavior. All of this information is localized and the implementation information is hidden. The value information defines what is known about the entity. The implementation of the value information, though hidden, is usually in terms of a data structure. The behavior of the entity is defined in terms of the possible operations upon the entity. The interface to the entity is defined in terms of these operations. Syntactically these operations are defined using function or procedure headers.

An "object" is like a variable except for the additional behavior information. Another aspect of object-based software development is the notion of "object-class". Just as "types" are used to define variables in Pascal, object-classes define objects in an object-based environment. An object defined by an object-class is said to be an instance of that class. The primary difference is that "typed" languages do not allow the programmer to restrict the set of associated operations to those defined by the programmer. In "typed" languages all operations inherited from the types used to define the new entity are also exported. Thus any variable of type "stack", where stack is represented by an array, can be accessed not only using programmer-defined stack operations but

also using array operations. Thus the designer cannot prevent someone from accessing or modifying the bottom element of a stack.

An object-class defines the set of legal values (called the domain of the object) for any instance of the class and it defines the behavior of any instance by specifying the set of legal operations that can be performed on an instance. The only means of manipulating an object is by the operations defined in that object's object-class. Associated with every object-class is a specification and an implementation. The specification, or interface definition, describes the means by which any other object in the system may interact with an instance of the object-class. The implementation, which is hidden from the user of an object-class, defines the data structure used to represent the domain and the implementation of the operations. Note that as long as the specification is not changed any changes to the implementation will not require modification to any other part of the system. A small example of a "list_of_students" object-class specification is given in figure 1.

The actual implementation is logically hidden from the user of the object-class. Almost any language can be used to code the implementation. Some languages such as Ada, Modula II, C++, and Smalltalk, prevent the user from accessing the hidden implementation. When using languages such as FORTRAN, COBOL, Pascal, and Basic the developers must all agree to only access an instance of an object-class by the means specified by its developer. These languages do not permit such agreements to be syntactically enforced. Enforcement usually comes during quality assurance inspections.

A "list_of_students" defined in this manner has several advantages. The object-class is highly independent of all other classes in the system, including the "student" object-class, as long as the operations in the "list_of_students" object-class are implemented correctly. The implementations for the operations in "list_of_students" may use knowledge of how the list is implemented but should access any other type of object only using operations defined and exported by that object's object-class. For instance, the operation 'sort_student_list' may be found in the "list_of_students" object-class. The implementation of the sort operation may require the comparison of two students to determine which comes first in the list alphabetically. Function-oriented developers using structured design for example may create the Ada code segment (shown in figure 2) to perform the needed comparison.

Unfortunately, much of the information concerning students and names is now embedded in the implementation of "list_of_students". This information includes: "student" is implemented as a record, the subcomponent of "student" that contains the name is called "name", and a name is a fixed length array. If the implementation of "student" or "name" changes for any reason then the implementation of "list_of_students" will need modification. A better object-based implementation is shown in figure 3.

List_of_Students

Domain: student* -- this means a sequence of 0 or more students

Operations:

Create:		->	List_of_students
Destroy:	List_of_students	->	
Add_Student:	List_of_students, Student	->	List_of_students
Delete_Student:	List_of_students, Student	->	List_of_students
Sort:	List_of_Students	->	List_of_students

Figure 1.
List_of_Students Object-Class.

The big difference between figures 2 and 3 is the independence of the various implementations. The "list_of_students" object-class need only have access to the 'less_than' operation of the "student" object_class. The "list_of_students" object_class does not know or need to know how "student" is implemented. The "student" object-class in turn, need only have access to the 'alphabetically_less_than' operation in the "name" object_class. Notice that the 'less_than' operation knows that a student is implemented as a record but does not know how a name is implemented. If we changed the implementations of any one of the object-classes, none of the other two would be affected. For instance, if the implementation of the "name" object-class were changed from a fixed-length array to a simple linked list then only the implementation of 'alphabetically_less_than' would be affected. Not only does this architecture make software easier to change, research has shown that it is much easier to understand and test.

Inheritance and Message Passing

Object-based software development can be thought of as the design of systems that have objects and object-classes where all information about a given object-class is localized and the implementation is hidden. Some object-based systems have extended this concept (Byte 1986). Such systems are called object-oriented systems and are based primarily on Smalltalk-like languages.

The two primary differences between object-based systems and object-oriented systems are the inclusion of the inheritance and message passing capabilities (Cox 1986). Inheritance allows the quick creation of new object-classes that are specializations of existing classes. All information common to both the specialization and the generalization is defined in the general "object-class". Only the additional information that distinguishes the specialization from the generalization is defined in the specialization object-class. For instance, assume that a "person" object-class exists. In Smalltalk-like languages a designer can create a new object-class "student" by stating that the object-class "student" is just like a "person" except for the addition of GPA and CLASS_SCHEDULE. The information and operations common to persons and students, such as NAME, ADDRESS, and 'less_than' are defined in the "person" object-class. The specialized

```
k:=1;
equal:=true;
scanned_both_names:=false;
WHILE equal AND NOT scanned_both_names LOOP
   IF (k > max_name_length) THEN
        scanned_both_names:=true;
        less_than:=false;
   ELSIF list(i).name(k) < list(j).name(k) THEN
        equal:=false;
        less_than:=true;
   ELSIF list(i).name(k) > list(j).name(k) THEN
        equal:=false;
        less_than:=false;
   ELSE
              k:=k+1;
   END IF;
END LOOP;
```

Figure 2.
Comparing Two Students.

information such as GPA, CLASS_SCHEDULE, and 'get_gpa' are defined in the "student" object-class. This ability provides several advantages. First, it is easier to create new object-classes without reinventing common aspects of both object-classes. Second, the software is easier to maintain since those domain and operation implementations common to several specializations exist in only one location (the generalized object-class). Because of inheritance the common implementations need only be implemented once, tested once, and if needed, modified once. In other words, the inheritance mechanism reduces redundant code and redundant creation and modification effort. The third advantage is conceptual. Inheritance allows generalization/specialization abstractions. This is a capability not provided by other languages.

```
IF less_than(list(i), list(j)) THEN
```

Determine if One Student is Less Than Another
Code Segment Found in 'Sort_Student_List'
An Operation in "List_of_Students"

```
less_than:= alphabetically_less_than(student_1.name, student_2.name);
```

Two Students are Compared Using Their Names
Part of the Implementation of 'Less_Than'
An Operation in "Student"

```
k:=1;
equal:=true;
scanned_both_names:=false;
WHILE equal AND NOT scanned_both_names LOOP
   IF (k > max_name_length) THEN
        scanned_both_names:=true;
        less_than:=false;
   ELSIF name_1(k) < name_2(k) THEN
        equal:=false;
        less_than:=true;
   ELSIF name_1(k) > name_2(k) THEN
        equal:=false;
        less_than:=false;
   ELSE
        k:=k+1;
   END IF;
END LOOP;
```

Definition of What it Means for One Name to be Less Than Another
Implementation of 'Alphabetically_Less_Than'
An Operation in "Name"

Figure 3.
Object-Based Version of Student Comparison.

The other extension to object-based systems is message passing. Message passing in object-oriented languages is not usually a concurrency concept (though it should be) but is the late binding of an operation invocation to an actual operation. Since each object has a domain and its own set of operations, Smalltalk-like languages make it possible to indicate the function to be executed relative to an object in question. This is done at execution time. For instance, assume that two objects, person A and desk B, each had a "print" routine associated with them. Furthermore, assume that variable X is untyped so that it can be assigned the value A or B. In object-oriented languages, operations similar to the following program segment are legal:

```
print(X).
```

In this case, the actual "print" routine called is determined by whether X has the value A (it is a person) or has the value B (it is a desk). When the program gets to the print statement it determines the type of X and calls the appropriate print procedure. In other languages this is not possible because the binding of a procedure to a procedure call is done at link time. This forces the programmer to explicitly code the decision-making process when trying to simulate late binding. The following segment demonstrates how the print operations might be done in Ada:

```
IF X.typ = person THEN
   person.print(X.person_info); ELSE
   desk.print(X.desk_info); END IF;.
```

In Ada and similar languages this can become difficult to understand and maintain if we want to add other types. In Smalltalk, no changes need to be made.

Late binding has advantages and disadvantages. It can help understandability and maintenance if used judiciously. It can also cause problems. First, it definitely slows down the program execution. Second, if X had been assigned the car C, then the print operation would only work if there is a print operation associated with cars. The programmer must make sure that all objects assigned to X have the operations "print". Most other languages are able to check at compile time to make sure all operations are available. Third, if the implementation of the "print" routine in the object-class "desk" doesn't print but actually erases all of the desk information then the "print(X)" operation may have disastrous results when X is a "desk".

Inheritance and message passing are valuable tools for the experienced software developer. However, most development and maintenance problems can be overcome using object-based information localization and information hiding. Fortunately, we do not need special languages to use these new concepts though object-based languages do help.

Implementation Languages

Object-based software has been implemented using almost every language available including FORTRAN, COBOL, and Pascal. The newer languages such as Ada, Modula II, C++ and Eiffel are easier to use because they provide syntactic representations (e.g., packages, modules, classes) of object-classes. These representations allow the programmer to put all information about an object-class in one location. The languages are also defined so that a user of the abstraction may only access the information or operations exported by the designer. All other access is prevented. The difference between the newer and older languages is similar to the difference between languages that provide procedures and the languages that do not (such as Basic and Assembly). Languages with procedures allow the programmer to put all of the procedure information in one location and prevent users of the procedure from accessing the implementation directly.

Even if the language does not syntactically provide for object abstraction it is often possible to achieve the same effect through other means. In FORTRAN, a subroutine can be treated as an object (not object-class) with its own operations. The operations are usually represented as multiple entry points in the subroutine. Normally the subroutine is not accessed by name but only through the entry points. The code executed from one entry point is made independent of all other entry points so that each entry point acts like a separate subroutine. In C and some versions of Pascal, such as Turbo Pascal, we can use include files and macro pre-processors to simulate object abstractions. Each include file contains the exported domain and operation interface information for an object-class. The implementation is stored in another file and linked into the program later. Any user of an object-class uses the include file defining the object-class. This mechanism does not always hide all information.

Object-based development can even be implemented using COBOL and Basic. In this case the design is done using object-based concepts and then translated into the given language. Since the syntax of the language won't prevent it, it is up to the programmer and quality assurance personnel to prevent unauthorized access to hidden information.

There is one major disadvantage to using languages not designed for object-based programming. The methodology often leads to operations being implemented as calls to calls to calls. For instance, if a "name" is implemented as a "string" which in turn is implemented as a "list_of_characters" then the implementation of 'print_a_name' may be a call to 'print_a_string' which in turn calls "print_a_list_of_characters". This can cause a serious degradation in performance. The compilers for languages designed to support object-based programming have optimizers to eliminate this problem. In other languages, if performance is a problem, then the code needs to be optimized by hand. If the unoptimized and hand-optimized versions are not maintained separately and concurrently then major maintenance problems can be introduced.

OBJECT-BASED DESIGN

The advantageous of object-based software development are built into software during software design. The primary difference between object-based design (often referred to in the literature as object-oriented design) and other design techniques is that decomposition of the design is done with respect to objects, not procedures or data.

The basic method has several steps:

(1) Using the specification, identify as many of the object-classes in the specification as possible. Often, each object-class will have only one instance in the specification. Also, few, if any, of the classes will have any specified operations. It is usually easiest to list only each entity in the specification. For instance, when designing an operating system some of the object-classes listed might be files, users, disks, memory, and terminals.

(2) Identify the operations in the specification and associate them with an object-class. For instance, the specification of the operating system may require the ability to create files. This operation would be associated with the "file" object-class identified in the first step. Sometimes, as we identify operations, we notice that we have left out an object. In that case just create it and add the new operation. For instance, the specification of the operating system may require a means for logging into an account. Since an "account" object-class had not been created in the first step it would at this time be added to the list of object-classes and the 'login' operation added to it.

(3) Design the implementation. Start by defining the domain for an object-class. This is usually done by defining the data structure representation. As this is done, new object-classes will

be discovered and should be added to the design. For instance, assume the domain of a "file" object-class is implemented as follows.

```
TYPE file = RECORD
    file_name:name;
    kind:file_type;
    contents:disk_space;
END;
```

From this definition we can identify the object-classes "name", "file_type", and "disk_space". If these object-classes are not in the design, they are then added.

(4) After defining the domain for an object, design as many of the operations for the object-class as possible. As the operations are implemented many existing object-classes will be employed, maybe defining new operations for those classes. New object-classes may also be needed implement the new operation. For instance, when designing the 'create_file' operation in the "file" object we may need an "access_privileges" object. At that time we simply create a new, partially-complete "access_privileges" object-class and continue designing the operation.

(5) When the definition for an object-class is finished, follow steps 3 and 4 for any class that is not yet finished. As the new class is completed new object-classes will probably need to be created and new operations added to it or existing object-classes. At first the number of new object-classes will explode but this will die out in time.

(6) After all of the object-classes have been created, create all instances that are needed to complete the design. Most of the instances have already been created as variables in implemented operations. Usually only global objects found in the specification need to be created.

(7) If needed, hand optimize the design. Be sure to keep both copies of the design and maintain them.

It should be noted that this is not a lock-step methodology. It is not required that one step be completely finished before going on to the next. It is also natural to skip from step to step. It is only necessary that the designer be sure that when done all objects and needed operations have been specified and designed.

GUIDELINES FOR OBJECT-BASED DESIGN

Any design should be easy to understand and modify. This is usually accomplished by making the units of abstraction as cohesive and independent of the rest of the system as possible. When dealing with structured design the units of abstraction are functions. Myers and Constantine have given us good guidelines, called coupling and cohesion guidelines, for creating independent functions (Myers 1978, Yourdon and Constantine 1979). These are very useful and should be used when developing the functions found in an object-class. The following coupling and cohesion definitions and guidelines have been developed to facilitate the development of better object-based designs (Embley 1987).

Cohesion

(1) An object-class has *separable* strength if it exports two or more object domains that are independent. A separable object-class should be rewritten as several separate object-classes, one for each independent object domain in the original object-class. For instance, if an object-class

defines both a person and a house and all the operations are associated with either a person or a house then two object-classes should be created, one for person and one for house.

(2) An object-class has *multifaceted* strength if it exports two or more object domains that are not independent. Usually the object-class contains an operator that is defined over more than one exported domain. For instance, assume an object-class defines a person and class_grades and there exists an operation in the object-class that updates the grades for a given student. Both the class_grades and student are passed in as parameters. The operation is implemented by computing the grade and storing it in the list of grades. Unfortunately, the implementation of the operation probably uses knowledge of the student and class grades implementation. This object-class, as well as all multifaceted object-classes, should be disentangled and reencapsulated as several object-classes. In this case both a "student" and "class_grades" object_class should be created. The "student" object-class should have a 'compute_grade' operation and the "class_grades" object_class should have a 'store' operation. This structure promotes the independence of students and class grades.

(3) An object-class has *non-delegation* strength if there exists an exported operator (other than a general store or retrieve operator) that operates on a proper subset of the components that comprise the definition of a compound exported domain. Operators that have non-delegation characteristics should be removed and delegated to an object-class in which all subcomponents are used. It may be necessary to create new object-classes to accommodate these operators. For instance, assume a person object-class has an operation that computes the month's pay for the person by multiplying the person's wage times the number of hours worked. The actual computation of a month's pay should be delegated to a "pay" object. If not, then when we wanted to change the method of computing pay, accounting for overtime for instance, we would change the "person" object-class and not the "pay" object-class.

(4) An object-class has *concealed* strength if it includes unencapsulated object-classes in its implementation. New independent object-classes should be formed from the unencapsulated object-classes. This failure to abstract appears in many forms. One, for instance, is to locally declare an age as integer. Age really isn't an integer but is a new type. It should be declared and encapsulated.

(5) An object-class has *model* strength if it does not have concealed, non-delegation, multifaceted, or separable strength characteristics. A model strength object-class has one and only one domain, and every operation applies to the one domain and should not be delegated to an object-class whose domain definition has fewer subcomponents. This is the best form of abstraction.

Coupling

Coupling levels are described from best to worst.

(1) Two object-classes have *nil* coupling if the two object-classes are totally independent. While this is desirable, a system is useless unless some object-classes are related.

(2) Two object-classes have *export* coupling if either an exported domain of one object-class is used in the other or an exported operation of one is invoked by the other. When two object-classes must be related, they should be related by export coupling.

(3) Two object-classes have *overt* coupling if one object-class invokes an operator of the other that has been (accidentally) exported as a side-effect of a domain definition. If possible, an object-class that permits overt coupling should be rewritten so that there can be no implicitly-exported

operations. This form of coupling usually occurs when the designer exports the representation of a domain so as to avoid rewriting some of the operations associated with the representation. For instance, assume a person implements a stack with the push, pop, and is_empty operations. Furthermore, the array with pointer implementation is made available to the user so the designer will not have to implement a top_of_stack or print_stack operations. It is assumed the user will use the corresponding array operations to do so. Unfortunately the user can also manipulate elements in the middle and beginning of the stack even if the designer did not intend for the user to be able to do so. If the stack representation is changed from an array to linked list disaster could follow. While it is impossible to prevent this type of coupling, in many languages that do not support object-based design it should be avoided. Unfortunately, in languages that can prevent such problems designers still explicitly export the domain implementation.

(4) Two object-classes have *covert* coupling if one object-class gains access to and explicitly uses non-exported information in the other. This situation usually occurs when a user of an object-class thinks that the rules apply only to others. This is a big problem in languages that do not support object-based information hiding. It can only be solved by strict enforcement of standards.

(5) Two object-classes have *surreptitious* coupling if one object-class shares a common assumption with the other object-class. This form of coupling is usually unintentional but very insidious. Though it is difficult, the assumptions should be detected and explicitly represented.

In general, systems should be designed to have minimal coupling (nil or export coupling) and model cohesion.

SUMMARY

It appears that object-based design will be the methodology of choice in the future. When considering maintenance, modification, understandability, and testing, object-based design is superior to any other existing methodology. It does take some work to learn how to use. For many, it is difficult to shift from a function-based view of the world to an object-based view. Such problems are similar to the problems encountered by those shifting from assembly language to high-level languages or those that shifted from unstructured to structured programming. Like these other paradigm shifts, switching to object-based design is not easy but it is well worth the effort. It is not free; object-based design reduces human effort by increasing the need for more computer resources. This has, however, been the way of progress in the past. Let machines do more so we can do less, to accomplish the same task.

REFERENCES

Booch, G. 1986. Object-oriented development. IEEE Transactions on Software Engineering. February 1986. SE-12:211-221.

Byte, vol. 11, no. 2, August 1986.

Cox, B.J. 1986. Object Oriented Programming. Addison-Wesley Publishing Co., Reading, MA.

Embley, David W., and Scott N. Woodfield. 1987. Cohesion and coupling for abstract data types. Proceedings of the 1987 Phoenix Conference on Computers and Communications. February 1987. Scottsdale, AZ.

Myers, G.J. 1978. Composite/Structured Design. Van Nostrand Reinhold Company, New York, NY.

Wegner, P. 1987. Dimensions of Object-Based Language Design. Sigplan Notices-Oopsla. 22:168-182.

Yourdon, E., and L.L. Constantine. 1979. Structured Design: Fundamentals of a Discipline of Computer Program and Systems Design. Prentice-Hall, Englewood Cliffs, NJ.

DISCUSSION OF THE PAPERS PRESENTED IN TECHNICAL SESSION 8, PART 1: SOFTWARE ENGINEERING

Daniel Hoggan[1], Presiding
Donna Falkenborg[2], Recorder

PAPERS DISCUSSED

The Potential Impact of Software Engineering on Water-Quality Modeling by D.G. Jamieson and K. Fedra

Object-Based Software Development by S.N. Woodfield

SPECIFIC QUESTIONS AND COMMENTS

Question: (D. Aum, Missouri) Would you give an illustration to help students better understand the object-oriented software concept?

Response: (S. Woodfield, Brigham Young University, Salt Lake City, Utah) An example is the designing of a system with multiple windows on a screen. I'm going to look at that screen. For instance, there are several windows on this screen. What is a window? Since I'm going to have many of them I want to design a class definition or a template to define windows. And for each window there are several attributes. Now I'm going to define the domain of the window. What are the things about a window I want to know? I want to know its height, its length, where it's positioned on the screen. I may even want to know the background color. I may wish to know what information is contained in the window. Then there are things I might want to do to the window. I may wish to create a new window, to move the window, to revise the window, delete the window. Having defined it in this way, other people in the system can now say, well I need a window for the menu. So I'm going to call the window template and, say, create a window that I can open, put a template in, close, resize, and use most of the capabilities associated with a window in the interface. In fact, the first major application of the object-oriented paradigm was the windowing interfaces on the MacIntosh and Xerox software. A lot of them used this object-oriented model to develop those interfaces. We have since started to apply that model elsewhere.

Question: (D. Green, University of Hawaii) How does one take the output from the simulation and link it to decision making, and how would the object-oriented software approach be applied?

Response: (S. Woodfield) Object-oriented design does not apply to translating simulation output into a report to be used in another agency. Object-oriented design is more "how do I build the simulation in the first place, or how do I write software that does the translation?" But it doesn't really help you do the translation.

Comment: (R. Knight, North Dakota State University) Please comment on what the hardware platform will be for some of these expert systems or object-based programs in the next 5 to 10 years.

[1]Daniel Hoggan, Professor, Utah Water Research Laboratory, Utah State University, Logan, Utah.
[2]Donna Falkenborg, Logan, Utah.

Response: (D. Jamieson, Thames Water Authority, United Kingdom) We will see the continuance of mainframe computing as the corporate resource, feeding sophisticated workstations. Nothing new from what already exists, just bigger, better, and less expensive.

Response: (S. Woodfield) The mainframe as a centralized machine seems to be giving way to mini-supercomputers -- still a centralized machine, but not a mainframe in a sense, but more of a dedicated, yet very powerful, mathematical processor. And at the same time, associated with it, will be several very powerful workstations with great capabilities with respect to data bases, user interfaces and graphics. In the next 5 years you'll see a trend toward a dedicated number-cruncher, but it will still be a small, almost personal, computer. And also distributed about it will be the workstations.

Question: (A. Lumb, USGS, Sterling, Virginia) In working with and developing some software we didn't know about object-oriented programming, but we may be doing some of it. What about the separation of the data aspects, the software aspects, and the data base aspects? You talk about a stream and you go get a bunch of data - how is the data base aspect tied into this?

Response: (S. Woodfield) When new ideas emerge, there usually is a group already doing it. I would expect simulation people to already be using it. Coupling data bases with an object-oriented paradigm is one of the research areas of object-oriented development. Most software developers still separate those concepts. The problem is, I have this program that runs on the computer, I have data that sits out on disks. The object-oriented paradigm applies to the program and we have no idea what to do with the data. Object-oriented data base people are trying to make it look as if they are using the same concept. It will take 5-10 years to integrate.

Question: (K. Seip, Center for Industrial Research, Norway) What is the difference between Small-Talk and Lisp?

Response: (S. Woodfield) Lisp originally was a very function-oriented system. Everything in that language is either a list of things or a function. Lisp is also a very strong language that can be adapted to your own conceptual model. It is possible to use concepts of an object-oriented system. You can think from an object-oriented framework and translate it into various programming models (Lisp, Ada, Small-Talk, Pascal, Fea, FORTRAN). Lisp is a language that lends itself to interpretation in almost any conceptual model. The translation can be made quite easily but learning a new language is difficult. With Small-Talk, there's a conceptual way of thinking there's an easy way to talk about reality. Instead of Small-Talk representing what would be nice to do, it represents things that are nice to implement. They haven't designed the language to make it nice to model reality. They really have implemented the language so it runs fast on a computer. Nice, but you still have to make some sort of conceptual switch or map to go from reality modeling to your language. The concepts of Small-Talk seem to be better than some 10-20 year old languages.

Question: (D. Gustafson, Monsanto, St. Louis, Missouri) I am an engineer who has used FORTRAN or Pascal for 10 years and needs to have sophisticated mathematical equations implemented in the code. Is it worthwhile to try to learn any of the new higher-level languages in order to implement object-oriented programming?

Response: (S. Woodfield) The first language I used in objective-oriented software was (in 1977) FORTRAN and I didn't find much difficulty mapping the paradigm onto FORTRAN. The paradigm works very nicely in your highly interactive interface-type situations. In structured mathematical situations you don't need these higher-level languages.

Question: (Audience) What would the nitrate concentration model be used for now that your problem is solved?

Response: (D. Jamieson) The model is not a live model at the present time. It was developed for a specific purpose. It met the requirements of that purpose. It enabled the deferral of heavy capital investments and has paid for itself many times.

Question: (J. Atwood, Soil Conservation Service, Iowa State University, Ames, Iowa) What are the relative costs of this sort of programming language? We do a lot of complex processing of data and generally have data stored on various tapes. We use a lot of job control language, and then we bring this in and submit it as a batch mode rather than an interactive mode. What we're finding is that it takes a lot of programming time to get this set up. But as compared to some of the newer data base management programs, computational costs are a lot smaller with our current technique than with these newer ones.

Response: (S. Woodfield) I assume you are referring to the data base aspect of a lot of modeling. Data bases in a batch model will always be cheaper than in an interactive mode. It depends on what you want to optimize. Do you want to optimize machine time or people time? If it's people time you probably ought to go with fourth-generation languages and not expect any results in your lifetime. If you want to go with faster processing there's batch processing that sometimes will be far more beneficial. However, it still appears machine costs will continue to decline. At some point it will be cheaper to move over to these new technologies. With masses of data you are probably better off with what you have, at least for the time being.

Question: (R. Hartzog, Louisiana Department of Environmental Quality, Baton Rouge, Louisiana) We are trying to use geographic information systems for several projects and one of the main problems is the vast amount of time that is involved in digitizing the maps. What about new and future ways to digitize maps and put this information in the data base?

Response: (D. Jamieson) The UK is better off than the United States because in the UK the base maps are being digitized and will be completed for the whole country by 1991. We have had urgent needs and we have digitized our own base maps, but this is the exception, not the rule.

APPLICATION OF PARAMETER ESTIMATION TECHNIQUES TO SOLUTE TRANSPORT STUDIES

Martinus Th. van Genuchten[1], Steven M. Gorelick[2] and William W-G. Yeh[3]

ABSTRACT

This paper briefly reviews the utility of optimization methods for estimating selected parameters affecting solute transport in soil and groundwater systems. Whereas parameter estimation techniques have been commonly used for the inverse problem involving saturated flow, their application to solute transport processes in saturated and partly unsaturated systems is relatively new and has only recently been explored. Nevertheless, a number of laboratory and field applications currently exist that clearly indicate their potential for improved designs and analyses of solute transport experiments. Several examples illustrate practical applications of parameter estimation methods to different transport problems. Advantages and limitations of parameter estimation techniques are discussed, and promising areas for further research and development outlined.

INTRODUCTION

Evidence is increasing that the earth's soil and water resources are being adversely affected by domestic, industrial and agricultural activities. The intentional and unintentional releases of chemicals into the environment have now gained the attention and concern of the public, largely through its potentially disastrous effect on both the immediate and long-term qualities of the environment. Chemicals migrating from municipal and industrial disposal sites, as well as radionuclides escaping from nuclear energy and waste storage facilities, percolate through the vadose zone and eventually may pose serious threats to the quality of ground and surface waters. Fertilizers and pesticides intentionally applied to agricultural lands inevitably also move below the root zone toward the groundwater table.

In efforts to better manage and monitor the environmental migration of these chemicals, scientists and engineers over the years have developed increasingly complex computer models which predict the fundamental processes of water flow and solute transport in porous media. For example, computer models are now routinely used in research to integrate the most pertinent physical, chemical and microbiological processes operative in the root and vadose zones of agricultural soils. Also, simulation models have become extremely valuable for evaluating alternative management practices for the purpose of increasing crop production and/or limiting groundwater pollution. This trend of using computer models as tools in research and management will probably accelerate in the future as computer costs decrease and the need for more realistic models increases.

Unfortunately, reliable application of computer models to field-scale flow and transport problems requires the quantification of a large number of model parameters. The required experimentation can be extremely time-consuming and costly,because field-scale processes exhibit considerable

[1]Martinus Th. van Genuchten, Research Leader, USDA-ARS, U.S. Salinity Laboratory, Riverside, CA.

[2]Steven M. Gorelick, Associate Professor, Department of Applied Earth Science, Stanford University, Stanford, CA.

[3]William W-G. Yeh, Professor, Civil Engineering Department, University of California, Los Angeles, CA.

variability in both time and space. As the conceptual understanding and numericalexpertise to simulate increasingly complex systems increases, the accuracy of future simulations hinges on the quality and completeness of experimental data. Thus, there is a great need for more efficient and accurate methods for determining parameters that appear in our computer models.

Among the most important parameters affecting flow and transport in soil and groundwater systems are the fluid flux, the solute diffusion and dispersion coefficients, and as a large number of parameters accounting for various physical, chemical and microbiological interactions in the soil-water system. Traditionally, these parameters have been determined by imposing rather restrictive initial and boundary conditions so that exact analytical or approximate numerical solutions of the governing equations could be inverted directly. These direct methods generally lead to specific functional forms of the parameters in terms of measurable soil properties. Klute (1986) gives a recent inventory of such methods applicable to a wide variety of soil physical processes operative in the unsaturated zone. Contrary to these direct methods, parameter estimation techniques do not pose inherent restraints on the mathematical form of the governing equations, on the initial and boundary conditions, or on the invoked constitutive relationships. Another advantage of estimation methods is their ability to provide information about parameter uncertainty, a feature which is rarely possible with direct inversion methods.

Although parameter optimization techniques have been popularly applied to groundwater flow problems (Cooley 1985, Yeh 1986), their application to transport studies has been limited. Most efforts thus far have been directed to one-dimensional contaminant transport models, with a few restricted two-dimensional studies. One of those was by Murty and Scott (1977) who developed an algorithm for estimating the longitudinal and transverse dispersivities in a two-dimensional aquifer solute transport. These authors used interpolation to develop concentration polynomials which were then used in an inverse procedure which minimized the difference between observed and calculated concentrations. The method was found to be very sensitive to noise in the observed concentration data. Umari et al. (1979) used an optimization approach to similarly estimate groundwater solute dispersivities in a two-dimensional transient transport model. Quasi-linearization was used to perform minimization which required the solution of a sequence of linear programming problems. The method required large computer storage since a numerical model was included in the constraint set.

Gorelick et al. (1983) later investigated the problem of pollutant source detection. Their work considered the estimation problem as one in which the source terms in the convection-dispersion equation were unknown. For cases in which the pollutant sources did not materially affect the water flow field (concentration sources), the problem could be treated as a linear simulation-regression problem. A method-of-characteristics two-dimensional numerical transport model (Konikow and Bredehoeft 1978) was included in regression formulations which "matched" simulated and measured concentrations. Through the governing differential equation, the source locations and magnitudes could be identified. One formulation employed unconstrained stepwise least-squares regression, while another used constrained minimum absolute deviation regression. The latter method was found to be more robust. Both methods gave confidence intervals indicating the likely presence and magnitude of multiple pollutant sources for steady-state and transient problems. The general method is valuable for problems in which the flow field is well-defined.

Nonlinear least-squares inversion techniques were used by van Genuchten (1980, 1981) and Parker and van Genuchten (1984) to estimate a variety of transport parameters from observed spatial and/or temporal concentration distributions. Their method was applied to both equilibrium and two-site/two-region type non-equilibrium one-dimensional transport models for which analytical solutions of the governing equations could be derived. Parker and van Genuchten (1984) also applied the inverse technique to a field-scale stochastic model which assumes a lognormally distributed pore water velocity. Jury and Sposito (1985) later used least-squares, maximum

likelihood and the method of moments to estimate the unknown parameters in several transport models, including the classical convection-dispersion equation (CDE) and a stochastic transfer function model developed by Jury (1982).

In a more refined approach, Wagner and Gorelick (1986) studied the problem of parameter estimation for equilibrium transport in soils described by the CDE model, and for nonequilibrium transport in a river system containing dead-end zones. While the optimization problem itself was solved with a quasi-Newton method, a post-optimization analysis was also carried out using Monte Carlo simulations to evaluate the reliability of the estimated parameters. One of the important conclusions reached by Wagner and Gorelick (1986) was that parameter estimates obtained with spatial concentration data generally were much more reliable than estimates generated with temporal data. This was true for most parameters involved (dispersion coefficients and decay coefficients), but not for the flow velocity. Somewhat similar conclusions were reached by Yeh and Wang (1987) who used a modified Gauss-Newton method for estimating the dispersivity in several one- and two-dimensional transport models. They found that if prior information on sampling design is not available, spatially distributed data should be used in preference to temporally distributed data for more reliable parameter estimates. Their study showed that data must be collected at those points, in time and space, which exhibit high parameter sensitivities. These points are generally located around the concentration front. This same conclusion also follows from the work by Knopman and Voss (1987, 1988) who calculated sensitivities of solute concentrations to changes in several transport parameters.

In a different approach, Strecker and Chu (1986) combined method-of-characteristics solute transport simulation (Konikow and Bredehoeft 1978) with quadratic programming to estimate dispersivities for a two-dimensional flow system. In this method all errors in estimating hydraulic conductivities (solution of a previous problem) are lumped into the estimated dispersivity values. Wagner and Gorelick (1987) extended earlier one-dimensional work to two dimensions and considered the problem of simultaneous estimation of groundwater flow and contaminant transport parameters. In a nonlinear simulation-regression procedure they "matched" both simulated hydraulic head and concentration values with those measured over space and time. Solution of this coupled inverse problem provided parameter estimates and covariances for the hydraulic conductivity, the effective porosity, and the longitudinal and transverse dispersivities. It was found that simultaneous estimation, which solves both the flow and transport equations, provides a great deal of information about parameters controlling contaminant transport. This last method employed finite-element simulation with a quasi- Newton algorithm developed by Dennis et al. (1981). Using a related simulation-optimization approach, Mishra and Parker (1988) recently found smaller estimation errors when they simultaneously estimated selected hydraulic and solute transport parameters from transient unsaturated flow and tracer experiments, as compared to first inverting the hydraulic data followed by inversion of the transport data.

In this paper we will briefly review the current status of parameter estimation methods and their advantages and limitations for determining the unknown parameters in several laboratory - and field-scale contaminant transport models. First, the classical transport equations are presented, followed by a general discussion of parameter estimation methods. A number of practical examples from the above literature are then discussed, thus illustrating the applicability of optimization techniques to different types of solute transport problems.

GOVERNING TRANSPORT EQUATIONS

Classical descriptions of solute transport in saturated-unsaturated soil/aquifer systems are usually based on the single-component convection-dispersion equation

$$\frac{\partial \theta c}{\partial t} + \frac{\partial \rho s}{\partial t} = \frac{\partial}{\partial x_i}\left(\theta D_{ij} \frac{\partial c}{\partial x_j} - q_i c\right) \quad [1]$$

where c and s = solute concentrations associated with the liquid and solid phases of the soil, respectively;
θ = volumetric water content,
ρ = soil bulk density,
D_{ij} = dispersion coefficient tensor,
x_i (i=1,2,3) = cartesian coordinates with $z=x_3$ taken positive downwards in this study,
t = time, and
q_i = volumetric fluid flux given by Darcy's law as

$$q_i = -K_{ij}(h) \frac{\partial h}{\partial x_j} + K_{i3}(h) \quad [2]$$

where h = soil water pressure head and
K_{ij} = hydraulic conductivity tensor.

For transient flow, q_i in equation 1 can be calculated from solutions of the unsaturated-saturated flow equation

$$\left[C(h) + \frac{\theta}{n} S_s\right] \frac{\partial h}{\partial t} = \frac{\partial}{\partial x_i}\left[K_{ij}(h) \frac{\partial h}{\partial x_j} - k_{i3}(h)\right] \quad [3]$$

where C = soil water capacity, approximated by the slope of the soil water retention curve, $\theta(h)$;
S_s = specific storage coefficient, and
n = porosity.

For ease of presentation we neglected in equations 1 and 3 all water and solute sources or sinks which may result from a variety of physical, chemical or biological processes (e.g., water uptake by plant roots, extraction wells, precipitation/dissolution reactions, microbial transformations). Where needed, appropriate terms accounting for these sources and sinks may be added to the above flow and transport equations.

Deterministic solutions of equations 1 and 3 have been popularly used in contaminant transport studies. Unfortunately, these solutions are subject to a number of difficulties that are not easily resolved (Sposito et al. 1986, Nielsen et al. 1986). Among these are the hysteretic nature of the soil hydraulic functions $\theta(h)$ and $K_{ij}(h)$ in the unsaturated zone, the effects of temperature and soil salinity on the hydraulic functions, the extremely nonlinear dependency of K_{ij} on the pressure head during unsaturated flow, the neglect of air flow, the assumption that Darcy's law is valid for structured (fractured) media, constancy of the fluid density, and the assumption that the soil matrix compressibility is constant. Accurate prediction of solute transport in the vadose zone is further complicated by soil heterogeneities manifested at various scales in soils and aquifers. Thus, it is important to realize that any model, however refined and complex, remains a simplification of actual processes. This implies that estimated parameter values always include the effects of processes or properties that have been neglected in the model. A good example is the dispersion coefficient (D_{ij}) which in concept reflects the combined macroscopic effects of diffusion and mechanical dispersion, the latter resulting from variations in local fluid velocities inside individual pores, and between pores of different sizes, shapes and directions. In practice, however, D_{ij} includes all of the solute-spreading mechanisms which may have been omitted from the governing transfer equations, such as transient flow, various nonequilibrium phenomena, nonlinear

sorption, and field-scale variability. Because of this inability to include all transport mechanisms in a model, most or all previously established relationships between the dispersion coefficient and certain macroscopic properties (notably the fluid flux; see Scheidegger 1960) are therefore essentially empirical. Moreover, these relationships depend on the type of transport model used, and on the spatial scale and heterogeneity of the field site to which the model is applied.

Thus far, parameter estimation methods have not been applied to the above general description of saturated-unsaturated water flow and solute transport. In this study we will focus exclusively on transport during steady-state flow in a homogeneous system. If, in addition, a one-dimensional vertical soil profile is considered, equation 1 reduces to

$$\frac{\partial c}{\partial t} + \frac{\rho}{\theta}\frac{\partial s}{\partial t} = D\frac{\partial^2 c}{\partial z^2} - v\frac{\partial c}{\partial z} \qquad [4]$$

where $v = q_3/\theta$ is the average pore water velocity, and
$D = D_{33}$.

The second term of equation 4 describes interactions between the chemical in solution and that sorbed by the solid phase. Sorption or exchange can be described with equilibrium or non-equilibrium equations; both approaches will be discussed below. We will also discuss a stochastic model that considers the effects of areal variations in hydraulic fluxes on field-scale solute transport.

PARAMETER ESTIMATION METHODS

Since parameter estimation methods have been discussed at length elsewhere (Yeh 1986, Kool et al. 1987), we give here only a brief and general review. A parameter estimation problem may be formulated as a least-squares minimization problem in which an objective function O(b) is minimized through adjustment of some vector $\{b\} = \{b_1, b_2, ..., b_m\}^T$ of unknown parameters:

$$o(b) = \{c^* - c(b)\}^T [W] \{c^* - c(b)\} \qquad [5]$$

where $\{c^*\} = \{c_1^*, c_2^*, ..., c_n^*\}^T$ is an observation vector whose elements in this study represent concentrations; [W] is a symmetric weighting matrix and $c(b) = \{c_1(b), c_2(b), ..., c_n(b)\}$ represents the predicted response for a given parameter vector {b} at selected values of z and t which are commensurable with the observations.

The objective is to find the parameter vector $\{b^f\}$ which minimizes equation 5, and hence results in the "best" fit between model-predicted values and observed data. Provided the least-squares problem has a unique solution, the final parameter vector $\{b^f\}$ is also the best estimate for {b} given certain statistical assumptions about the underlying errors relative to the model.

Because measurements $\{c^*\}$ and model predictions {c(b)} are both subject to uncertainty, the parameter estimation problem is essentially a statistical problem. Maximum likelihood considerations lead to the choice of the inverse of the error covariance matrix as the weighting matrix [W]. In other words, [W] contains information about measurement accuracy and possible correlations between measurements. An obvious problem is the fact that often little or no information about [W] is available. This problem is usually resolved by making some a priori assumption about the structure of [W]. The greatest simplification arises when [W] is taken to be the identity matrix of dimension n, in which case equation 5 reduces to an ordinary least-squares objective function. This approach is equivalent to assuming independent and constant variance errors. A less restrictive formulation assumes uncorrelated but not necessarily equal variance

errors. [W] then becomes a diagonal matrix. This approach is appropriate when the observations $\{c^*\}$ consist of different quantities (e.g., pressure heads and concentrations) measured with different instruments and expressed in disparate units. In case of autoregressive error correlation, [W] can be expressed in terms of the autocorrelation coefficient which becomes an additional parameter to be estimated along with the elements of {b} (Beck and Arnold 1977).

Unless the model fitted is linear in all parameters, equation 5 must be solved iteratively. A variety of nonlinear optimization algorithms are now routinely available through software libraries. Gauss-Newton methods, and such modifications as the Levenberg-Marquardt algorithm, are typically used for solving least-squares type problems. Comparisons of different methods (e.g., Cooley 1985, Hiebert 1981) show that results can be very problem-dependent. When a parameter estimation problem becomes computationally more expensive to solve, it may become worthwhile to invest some time searching for the most efficient solution method for a specific application.

Inherent in the parameter estimation approach is the fact that one must assume some specific parametric form for a constitutive relationship (e.g., the dispersion coefficient as a function of the pore water velocity), which is presumed to hold to a sufficient degree of approximation. Clearly, if the adopted model does not represent the actual behavior of a system, then results of the optimization analysis will be of dubious value. Another source of parameter uncertainty is the ill-posedness of many inverse problems. Ill-posed inverse problems usually are characterized by nonuniqueness and instability of the parameter solution (Yeh 1986). The uniqueness problem has great practical importance. In the case of nonuniqueness the estimated parameter values will differ with changing initial estimates of the parameters.

As a consequence, model predictions may differ from inputs that deviate from those used in the parameter estimation process. Chavent (1974) concluded that, in general, the inverse problem of parameter estimation in a distributed system is inherently nonunique due to lack of adequate observations. As pointed out by Yeh (1986), the uniqueness problem is closely related to the notion of parameter identifiability, i.e., whether or not accurate estimates of the parameters in a given mathematical model can be obtained from available data. Identifiability thus depends crucially on the available experimental data. Instability occurs when the estimated parameters are excessively sensitive to small changes in the observed data. Relatively small measurement errors can then lead to significant errors in the estimated parameters. Proper model selection and sampling design, as well as accuracy of measurement, are often crucial to success of the parameter estimation problem; they should be given serious consideration prior to conducting experiments.

RELIABILITY OF PARAMETER ESTIMATES

Statistical measures can be used for analyzing the goodness of fit, the reliability and the correlation of the estimated parameters. It is important to note that these measures are strictly correct only for linear regression models, and that they are approximations for the nonlinear case. Let us also assume, without losing generality, that the weighting matrix [W] is equal to the identity matrix.

The variance of the residuals, σ^2, can be computed by

$$\sigma^2 = \frac{SSQ(\bar{b})}{M\text{-}L} \qquad [6]$$

where $SSQ(\bar{b})$ = the sum of squares of residuals by using parameter estimate $\{\bar{b}\}$ with (M-L) degrees of freedom,
M = the total number of observations, and
L = the dimension of the parameter vector $\{\bar{b}\}$.

The variance of the residuals gives a good measure of the overall goodness of fit of the estimated model.

The covariance matrix of the estimated parameters in nonlinear regression can be approximated by the following form (Yeh, 1986):

$$Cov(\bar{b}) = \frac{SSQ(\bar{b})}{M-L} [A(\bar{b})]^{-1} \quad [7]$$

where $[A] = [J_D^T J_D]$, and
$[J_D]$ = the Jacobian matrix of c with respect to the parameter {b}.

In fact, the elements of the Jacobian are the sensitivity coefficients $\partial c/\partial b_j$.

The diagonal elements of $Cov(\bar{b})$ are the variances of the estimated parameters. A norm of the covariance matrix can be used to represent parameter uncertainty. Such norms as trace and determinant have been commonly used. Well-estimated parameters are generally characterized by small variances, as compared to insensitive parameters which are usually associated with large variances. By definition, the correlation matrix of the estimated parameters is

$$[\bar{R}] = \begin{vmatrix} \frac{r_{11}}{\sqrt{r_{11}r_{11}}} & \cdots & \frac{r_{1L}}{\sqrt{r_{11}r_{LL}}} \\ \frac{r_{L1}}{\sqrt{r_{LL}r_{11}}} & \cdots & \frac{r_{LL}}{\sqrt{r_{LL}r_{LL}}} \end{vmatrix} \quad [8]$$

where the r_{ij}'s are elements of the covariance matrix of the estimated parameter.

The more sensitive the parameter, the more closed and quicker the parameter will converge. A correlation analysis of the estimated parameters indicates the degree of interdependence among the parameters with respect to the objective function. Correlation of parameters is called the collinearity problem. Such a problem can cause slow convergence in minimization, and in most cases will result in non-optimal parameter estimates.

EQUILIBRIUM TRANSPORT

We first consider the simple case of equilibrium transport through a relatively short, homogeneously packed laboratory soil column. Adsorption or exchange reactions perceived as instantaneous are described by equilibrium isotherms, s(c), which can be of the mass action, linear, Freundlich or Langmuir type, or of many other functional forms. Here we assume that sorption can be described by a linear isotherm of the form

$$s = kc \quad [9]$$

where k is a distribution coefficient which is independent of concentration. Substitution into equation 4 leads then to the classical convection-dispersion equation (CDE)

$$R \frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial z^2} - v \frac{\partial c}{\partial z} \quad [10]$$

where the retardation factor R is given by

$$R = 1 + \rho k/\theta. \quad [11]$$

For solutes exhibiting linear sorption, k is positive and R becomes greater than one. If there are no interactions between the chemical and the soil (k=0), R equals one. In some cases R may become less than one, indicating that only a fraction of the liquid phase participates in the transport process. This is the case when the chemical is subject to anion exclusion or when relatively immobile liquid regions are present (e.g., inside dense aggregates) which do not contribute to convective transport.

An example, equation 1 was used by van Genuchten (1980) to analyze the transport of the anion chloride through a 30-cm long soil column of repacked Norge loam. Figure 1 shows a plot of the measured column effluent data, together with a fitted curve obtained by nonlinear regression using the Levenberg-Marquardt method (Marquardt 1963). Analysis of the data was facilitated by casting equation 10 in reduced form using the dimensionless variables:

$$T = vt/L \qquad Z=z/L \qquad [12a,b]$$

$$P = vL/D \qquad [12c]$$

where T = the amount of pore volumes leached through the column,
x = reduced distance,
L = column length, and
P = the column Peclet number.

Introducing equations 12a,b,c into 10 gives

$$R\,\frac{\partial c}{\partial T} = \frac{1}{P}\,\frac{\partial^2 c}{\partial z^2} - \frac{\partial c}{\partial Z}. \qquad [13]$$

We also assume that concentrations have already been normalized with respect to the influent concentration, C_0. The displacement experiment involves the pulse-type application of a chloride tracer solution to an initially solute-free soil column. Applicable initial and boundary conditions for the experiment are then

$$c(Z,0) = 0 \qquad [14a]$$

$$\left(-\frac{1}{P}\,\frac{\partial c}{\partial Z} + c\right)\Big|_{Z=0} = \begin{cases} 1 & 0 < T < T_0 \\ 0 & T > T_0 \end{cases} \qquad [14b]$$

$$\frac{\partial c}{\partial z}\,(\infty,t) = 0 \qquad [14c]$$

where $T_0 = vt_0 / L$ is the number of pore volumes of tracer solution leached through the column (t_0 is the actual duration of the applied pulse).

Concentrations in the equations above are assumed to represent "volume-averaged" or "resident," values (Kreft and Zuber 1978, Parker and van Genuchten 1984 a or b). As shown by van Genuchten and Parker (1984), equation 14 closely approximates transport through a finite column as long as effluent concentrations are interpreted as "flux-averaged," concentrations. The analytical solution for the effluent curve $c_e(T)$ is then

$$c_e(T) = A(T) - A(T-T_0) \qquad [15a]$$

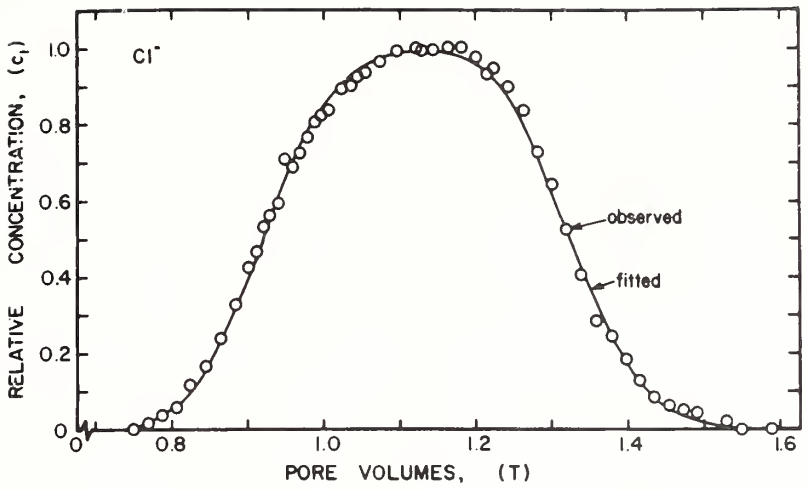


Figure 1.
Observed and fitted effluent curves for chloride transport through Norge loam.

where

$$A(T) - \frac{1}{2}\,\mathrm{erfc}\left[\left(\frac{P}{4RT}\right)^{1/2}(R-T)\right] + \frac{1}{2}\,e^{P}\,\mathrm{erfc}\left[\left(\frac{P}{4RT}\right)^{1/2}(R+T)\right] \qquad [15b]$$

Equations 13 and 15 reveal three independent parameters, P, R and T_0, while the original transport problem contains four parameters: v, D, R and t_0. One of the three coefficients v, D and R in equation 10 must be known independently. This follows immediately by noting that division of 10 by a constant permits one of the coefficients to be eliminated. In practice either v or R (or both) must be known beforehand, allowing D to be estimated from observed data. For adsorbing chemicals, R can at least in principle be determined from batch equilibration studies. Here we elect to use the dimensionless formulation and implicitly assume that v can be measured independently.

Figure 1 shows an excellent match of the fitted curve with the experimental data. The estimated dimensionless parameters and their standard errors of estimation are listed in table 1 (Example 1). Note that R is less than one, indicating some anion exclusion. Using equation 11 and independently measured values for θ and ρ (table 1), a value of 0.019 cm^3/g can be derived for the specific anion exclusion volume (-k).

The close fit of the calculated curve with the observed data in figure 1 was expected for the very homogeneous, sieved and repacked soil used in this experiment. Unfortunately, such close fits are not always possible, especially for aggregated soils and/or for chemicals that are strongly adsorbed. Figure 2A shows results for a tritiated water ($^3H_20$) effluent curve from a laboratory column packed with Glendale clay loam consisting of aggregates of less than 6 mm in diameter (van Genuchten 1981). Note the small but conceptually quite significant deviations between the fitted curve and the observed data at the higher concentrations and the higher pore volumes. Deviations of this type are hypothesized to result from physical and/or chemical non-equilibrium conditions during transport.

Table 1.
Fitted and measured parameter values and their standard error of estimation (±) for several transport problems.

Example:	1	2A	2B	3A	3B
Tracer:	Cl^-	^3H_2O	^3H_2O	2,4,5-T	2,4,5-T
	----------Fitted Values----------				
P	287. ± 12.	9.43 ± .84	30.7 ± 2.0	3.67 ± .40	25.9 ± 3.9
R	.922 ± .001	.974 ± .017	1.048 ± .006	2.223*	2.223*
β	-	-	.719 ± .008	-	.605 ± .015
ω	-	-	.534 ± .038	-	.495 ± .055
T_0	.406 ± .002	2.079 ± .019	2.103 ± .004	-	-
	----------Measured Values----------				
ρ	1.527	1.126	1.126	1.309	1.309
θ	.363	.401	.401	.456	.456
q	5.160	16.58	16.58	16.81	16.81
T_0	.425	2.08	2.08	4.95	4.95

*derived from batch equilibration data (fixed on input)

NONEQUILIBRIUM TRANSPORT

Many chemically controlled or diffusion-controlled rate reactions have been proposed over the years. The most popular and simplest expression arises when sorption is assumed to be a linear, first-order reversible process. The term $\partial s/\partial t$ in equation 4 becomes

$$\frac{\partial s}{\partial t} = \alpha(kc - s) \qquad [16]$$

where α is a first-order rate coefficient. Although equation 16 and similar rate laws have resulted in some improvements over equilibrium methods in predictive capability, success has generally been limited to experiments carried out at relatively low flow velocities.

A chemical nonequilibrium model that has led to improved predictions is the two-site model in which sorption is assumed to consist of two components, one governed by equilibrium and one by first-order kinetics (Selim et al. 1976, Cameron and Klute 1977). Basic to this model is the

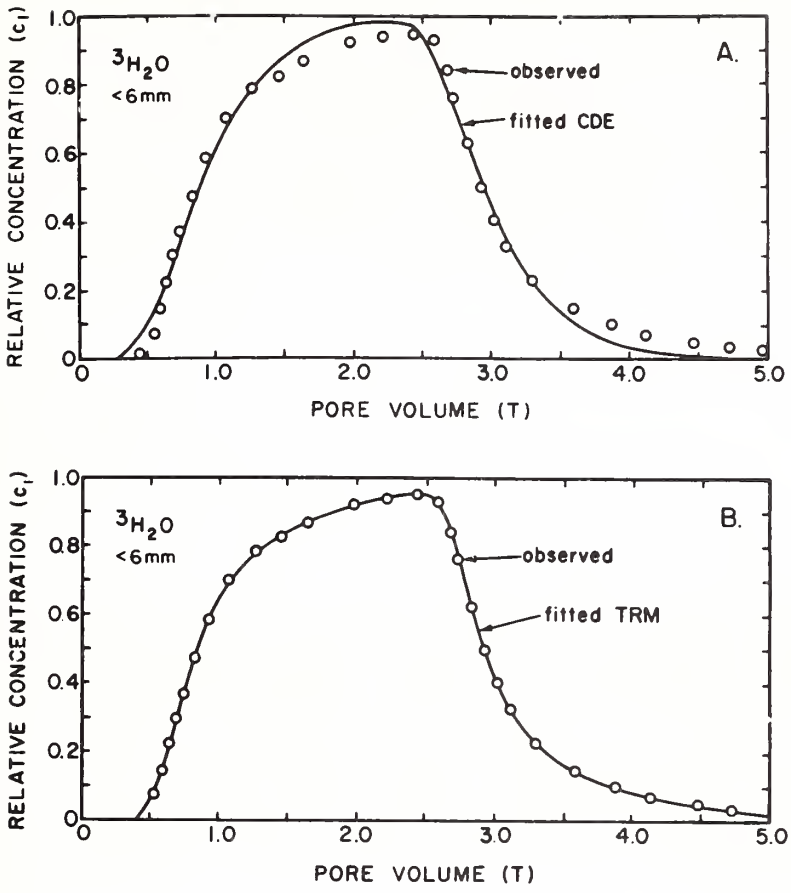


Figure 2.
Observed and fitted effluent curves for tritiated water movement through Glendale clay loam. Fitted curves were based on: (A) the classical CDE, and (B) the two-site/two-region TRM transport models.

assumption that chemicals will react with different constituents (soil minerals, organic matter, iron and aluminum oxides) at different rates and different intensities. The model divides the sorption sites into two fractions: equilibrium-controlled "type-1" and kinetically controlled "type-2" sorption sites. This conceptualization leads to the following model

$$(1 + \frac{f\rho k}{\theta}) \frac{\partial c}{\partial t} + \frac{\rho}{\theta} \frac{\partial s_2}{\partial t} = D \frac{\partial^2 c}{\partial z^2} - v \frac{\partial c}{\partial z} \quad [17a]$$

$$\frac{\partial s_2}{\partial t} = \alpha[(1-f)kc - s_2] \quad [17b]$$

where f is the mass fraction of all sites occupied with type-1 equilibrium sites, and where the subscript 2 refers to type-2 sorption sites.

Note that the one-site model (equations 4 and 16) is the special case of the two-site model when f=O. The two-site model can be expressed in the following dimensionless form (Nkedi-Kizza et al. 1984)

$$\beta R \frac{\partial c_1}{\partial T} + (1-\beta)R \frac{\partial c_2}{\partial T} = \frac{1}{P} \frac{\partial^2 c_1}{\partial z^2} - \frac{\partial c_1}{\partial z} \quad [18a]$$

$$(1-\beta)R \frac{\partial c_2}{\partial T} = \omega(c_1 - c_2) \quad [18b]$$

where the parameters R and P are the same as before, and where

$$\beta = \frac{\theta + f\rho k}{\theta + \rho k} \qquad \omega = \alpha(1\text{-}f)\rho kL/q \qquad [19a,b]$$

$$c_1 = c \qquad c_2 = \frac{s_2}{(1\text{-}f)k}. \qquad [20a,b]$$

Nonequilibrium has also been explained by assuming diffusion-controlled sorption. In many soils, especially when aggregated or containing various types of macropores, the sorption rate may be limited by the rate at which solutes diffuse to the sorption or exchange sites. This alternative viewpoint has led to physical nonequilibrium models that partition the liquid phase in mobile (flowing) and stagnant (immobile) phases. The approach in effect assumes a bimodal pore-water velocity distribution: convective-dispersive transport is limited to only a fraction of the liquid-filled pores, while the remainder of the pores have stagnant water. This stagnant or immobile water has been visualized as thin liquid films around soil particles, as dead-end pores (Coats and Smith 1964), as nonmoving intra-aggregate water (Passioura 1971), or as relatively isolated regions associated with unsaturated flow (Nielsen and Biggar 1961). Assuming first-order exchange of solute between mobile and immobile regions, this "two-region" model leads to (van Genuchten and Cleary 1979)

$$\theta_m R_m \frac{\partial c_m}{\partial t} + \theta_i R_i \frac{\partial c_i}{\partial t} = \theta_m D_m \frac{\partial^2 c_m}{\partial z^2} - \theta_m v_m \frac{\partial c_m}{\partial z} \qquad [21a]$$

$$\theta_i R_i \frac{\partial c_i}{\partial t} = \alpha(c_m - c_i) \qquad [21b]$$

where the subscripts m and i refer to mobile and immobile regions, respectively, and where α is a mass transfer coefficient, interpreted as a diffusion coefficient divided by some average diffusion path length. The retardation factors R_m and R_i account for equilibrium-type sorption processes in the mobile and immobile regions, respectively.

Comparison of the two-site (equations 17a, b) and two-region (equations 21a, b) non-equilibrium models shows that they have the same mathematical structure. The two-region model can be expressed in the same dimensionless form as previously used for the two-site model (equations 20a, b), provided the following model-specific parameters are used (Nkedi-Kizza et al. 1984, Sposito et al. 1986)

$$P = v_m L/D_m \qquad R = 1 + \rho k/\theta \qquad [22a,b]$$

$$\omega = \alpha L/q \qquad \beta = \theta_m R_m/R \qquad [23a,b]$$

$$c_1 = c_m \qquad c_2 = c_i \qquad [24a,b]$$

where $v_m = q/\theta_m$. Note that only R remains the same as before.

Because the same dimensionless transport equations (equations 18a,b) apply to conceptually different transport models, it follows immediately that effluent curves from laboratory columns by themselves contain insufficient information to differentiate between specific physical and chemical processes leading to nonequilibrium. Hence, independent parameter estimates are needed to effectively differentiate between the presumed two-site and two-region type nonequilibrium phenomena.

Figure 2B compares the same ^3H_2O effluent data as before with a fitted curve obtained with the analytical solution (Parker and van Genuchten 1984a) of the above two-site/two-region (TRM) model. While agreement with the observed data is now excellent, the improvement is obtained at the expense of two additional adjustable parameters. The fitted independent parameters P, R, β and ω are listed in table 1 (Example 2B). Note that R is now slightly larger than 1, indicating some isotopic exchange. The extent of this exchange, however, is so small, in this case, that it could have been neglected. Fixing R at one and T_0 at its measured value (see table 1) resulted in nearly the same fitted values for P, β and ω, and to essentially the same calculated curve.

Figure 3 shows another application of the classical (CDE) and two-site/two-region (TRM) models, in this case to transport of the pesticide 2,4,5-T (2,4,5-Trichlorophenoxyacetic acid) through 30-cm long columns packed with the same aggregated Glendale clay loam soil as before. Because of observed hysteresis in the adsorption-desorption isotherms only the breakthrough side of the curve was used in the optimization process. In this case three parameters were estimated (P, β and ω), thereby assuming that R was known. The batch equilibrium adsorption isotherm for 2,4,5-T was described well with the nonlinear Freundlich isotherm $s = 0.616c^N$ with N=0.792. Linearization of this isotherm gave a value of 0.426 for k (van Genuchten 1981) which in turn results in R=2.223 using equation 8 and the measured values for ρ and θ (table 1, Example 3). Figure 3B shows an excellent fit with the data. A one-parameter fit with the CDE model largely failed (Fig. 3A). Keeping R as an additional adjustable parameter in this model only marginally improved the modelled description of the effluent data (results not shown here).

Table 1 (Examples 2 and 3) indicates that the classical CDE equation generates much smaller values for P (and hence larger D-values) than the two-site/two-region TRM model when fitted to the same data. This is because the TRM model explicitly accounts for observed nonequilibrium while the CDE equation can include those effects only by adjusting the value of the dispersion coefficient. Similar effects on D also occur when nonlinear sorption/exchange is neglected. For example, nonlinear sorption causes sharpening of a concentration front when N<1, thus having the opposite effect of dispersion (van Genuchten and Cleary 1979). The interactive effects of dispersion, nonlinear sorption or exchange, chemical hysteresis, kinetic mechanisms and intra-aggregate diffusion on transport are discussed also by Kool et al. (1987), Jardine et al. (1985) and Parker and Valocchi (1986), among others.

The mathematical similarly of the two-site and two-region models suggests that the two models can be used to describe macroscopic transport behavior without having to delineate the exact physical and chemical processes at the microscopic level. A rigorous analysis of diffusion to less accessible sites should have been described by Fick's law of diffusion. This may be possible when the shapes and sizes of soil aggregates are known (Parker and Valocchi 1986, van Genuchten and Dalton 1986). Unfortunately, this is not easily done for a soil that contains a mixture of irregularly-shaped, small-sized aggregates; and ironically also not for seemingly homogeneous soils. Because of the fuzzy geometric distribution of immobile water pockets and associated sorption sites, and given our inability to microscopically measure these sites, several parameters in (18a,b), notably β and ω, often must be fitted to observed effluent or other data before the model can be used. While parameter optimization techniques such as those used in this study are important tools for that purpose, their use illustrates the semi-empirical nature of present models.

The examples above deal with carefully controlled laboratory experiments resulting in well-defined effluent curves and relatively small experimental errors. Unfortunately, experimental conditions are generally far less well-controlled in field studies, thus causing a variety of measurement and sampling errors. The effect of measurement error and other causes of variability in the data was recently studied by Wagner and Gorelick (1986). Their results as well as those of Jury and Sposito (1985) indicate that spatially distributed observations generally lead to better estimates and lower bias than observations distributed in time. Similar conclusions were also reached by

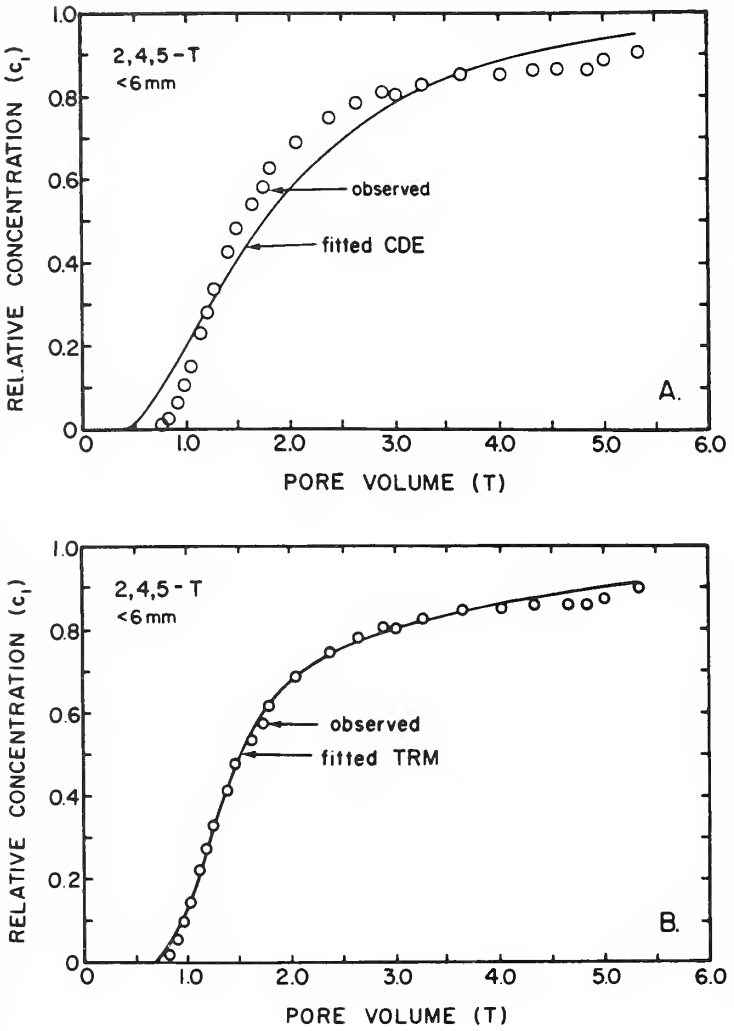


Figure 3.
Observed and fitted effluent curves for 2,4,5-T movement through Glendale clay loam. The fitted curves were based on the classical CDE (A) and two-site/two-region TRM (B) transport models.

Yeh and Wang (1987) and Knopman and Voss (1987, 1988) by calculating sensitivities of solute concentrations to changes in the parameter values.

FIELD-SCALE TRANSPORT

To evaluate the effects of field-scale variability on transport, a stochastic approach must generally be taken. We will limit our discussion here to a relatively simple formulation stemming from the studies by Amoozegard-Fard et al. (1982) and Parker and van Genuchten (1984a). The model is very similar to the one-dimensional models of Bresler and Dagan (1981), and Simmons (1982). Conceptually, the approach assumes that the entire field (referred to as the global scale) is composed of numerous independent parallel vertical soil columns, with transport in each soil column (termed the local scale) described by the CDE equation using constant coefficients. Lateral flow, transverse dispersion and vertical heterogeneities are ignored.

To approximate transient flow, hydraulic fluxes are averaged over the time and space domains of interest. At the local scale, mean local hydraulic fluxes $\bar{q}_0$ and water contents $\bar{\theta}$ are defined as (Parker and van Genuchten 1984a).

$$\bar{q}_0 = \frac{1}{t_m} \int_0^{t_m} q_0(t)\, dt \qquad [25a]$$

$$\bar{\theta} = \frac{1}{z_m t_m} \int_0^{t_m} \int_0^{z_m} \theta(z,t)\, dz\, dt \qquad [25b]$$

where $(0,t_m)$ and $(0,z_m)$ = the time and distance intervals for averaging, and
q_0 = the hydraulic flux at the soil surface.

The time-averaged local velocity is then taken as

$$\bar{v} = \bar{q}_0 / \bar{\theta} . \qquad [26]$$

At the global- or field-scale, the mean over the areal domain A of the time-averaged velocity $\langle \bar{v} \rangle$ is given by

$$\langle \bar{v} \rangle = \frac{1}{A} \int_A \bar{v}\, dA. \qquad [27]$$

Similar areal averages can be defined for the time-averaged surface flux $\langle \bar{q}_0 \rangle$ and the instantaneous surface flux $\langle q_0(t) \rangle$. The equivalent steady-state time variable is now defined by the transformation

$$t^*(t) = \int_0^t \langle q_0 \rangle (\tau) d\tau / \langle \bar{q}_0 \rangle . \qquad [28]$$

The local pore-water velocity $\bar{v}$ is assumed to vary lognormally among the different columns according to the probability density function $\rho(\bar{v})$:

$$\rho(\bar{v}) = \frac{1}{\bar{v}\sigma_{ln}\sqrt{2\pi}} \exp\left\{ - \frac{[\ln(\bar{v}) - \mu_{ln}]^2}{2\sigma_{ln}^2} \right\} \qquad [29]$$

where μ_{ln} and σ_{ln}^2 are the mean and variance of $\ln(\bar{v})$, respectively. The local dispersion coefficient is assumed to be perfectly correlated with $\bar{v}$ such that $D = \epsilon \bar{v}$ with a deterministic apparent dispersivity, ϵ. Fixing R at unity leads then to a three-parameter field-scale transport model described by ϵ, σ_{ln} and μ_{ln}.

Field-scale resident concentrations are given as areal means over the domain A as

$$\hat{c}(z,t) = \int_A c(z,t) dA \Big/ \int_A dA \qquad [30]$$

where c is the local concentration. Since $\bar{v}$ is the only random variable on A, $\rho(\bar{v})d\bar{v}$ may be substituted for dA to give

$$\hat{c}(z,t) = \int_0^\infty c(z,t^*;\bar{v})\, \rho(\bar{v}) d\bar{v} \Big/ \int_0^\infty \rho(\bar{v}) d\bar{v}. \qquad [31]$$

The integrals in this equation may be evaluated by numerical quadrature once local soil surface boundary conditions are formulated. As appropriate, these may be defined by imposing a constant pulse duration t_0 or a constant mass loading, the latter leading to local values of t_0 that vary inversely with $\bar{v}$.

We previously applied the above regional stochastic model (RSM) to a field tracer experiment (Jury et al. 1982) with bromide on a 0.64-ha field during transient flow (Parker and van Genuchten 1984a). Results are summarized in figure 4, which shows plots of the areally averaged resident concentration versus transformed time t^*, at various depths. Concentrations at 30 cm were used as input in the inverse problem to estimate ϵ, μ_{ln} and σ_{ln}. The fitted values in table 2 show large uncertainty in ϵ. A two-parameter analysis of the data with ϵ fixed at various values (see trials 2-5 in table 2) indicates little sensitivity of the model to ϵ for $\epsilon < 10$ mm. The mean error increases slightly when the dispersivity increases; some compensation between ϵ and σ_{ln} occurs as the value for σ_{ln} decreases for larger ϵ. Thus, field-scale dispersion is reduced as local-scale dispersion increases. Table 2 indicates that even when $\sigma_{ln}=0$ (which corresponds to the deterministic CDE model), the sum of squared deviations (SSQ) between observed and fitted concentrations is only moderately larger than for the three-parameter RSM model.

Fitted curves using trial 1 and trial 6 parameter estimates (see table 2), corresponding to the three-parameter RSM and two-parameter CDE models, respectively, are compared to observed data for the z=30 cm data in figure 4 (top). Breakthrough curves for the 60- and 90-cm depths, predicted with these estimated parameters, compare reasonably well with the observed data (middle and bottom). Figure 4 shows that even though the two models are virtually indistinguishable at the calibration depth, they yield increasingly divergent predictions at greater depths. Jury and Sposito (1985) presented a detailed comparison of the CDE model and a transfer function model (TFM; Jury 1982) which is similar to the RSM model described above with $\epsilon=0$ (no local-scale dispersion). Results of their calibration of the CDE and RSM models are shown in figure 5. Error ellipses computed by partitioning squared deviations between model predictions and observations by the chi^2 theorem indicate better clustering for the TFM than the CDE model, although both models appear too simplistic to fully describe the observed bromide transport data. The CDE model tends to underpredict the rate of spreading of the solute distribution due to the assumption of homogeneity, while the TFM tends to overpredict spreading due to the assumption of zero lateral transport and of perfect correlation between velocity distributions with depth. The results emphasize the caution one must exercise when a calibrated model is used to extrapolate to conditions, particularly to travel distances, which differ greatly from those used during calibration.

RIVER POLLUTANT TRANSPORT

We consider one more example involving parameter estimation in a convective-dispersive system, but now for surface water flow. This field application deals with a tracer experiment conducted on a mountain stream in Northern California. The experimental data and initial modeling were presented by Bencala and Walters (1983), whereas the parameter estimation work was carried out by Wagner and Gorelick (1986). In the experiment, chloride was added for three hours to Uvas Creek and concentrations measured over a 24-hour period at 5 locations extending 619 m downstream. Because of the long tailing in the breakthrough curves, a simple convective-dispersive model could not adequately fit the data. Bencala and Walters (1983) recognized the need for a transient storage submodel as part of the overall conceptual model of solute transport. The river system is divided into two transport zones. One is for the flowing stream in which the tracer moves by convection and dispersion. The other consists of the bed and banks of the stream channel which temporarily detain some of the tracer, eventually releasing it back to the stream. Movement in and out of this delayed-storage zone leads to a coupled model represented by the following equations:

Table 2.
Parameter estimates and their standard error of estimation (±) for the field-scale bromide tracer experiment of Jury et al. (1982). Values fixed on input are shown in parenthesis.

Trial	e (mm)	⟨ v⟩ (mm/day)	σ_{ln}	SSQ
1	1.0 ± 227.	30.5 ± 21.0	.800 ± .943	.0005177
2	(.01)	30.5 ± 1.8	.803 ± .060	.0005185
3	(1.0)	30.5 ± 1.8	.800 ± .060	.0005177
4	(10.0)	29.7 ± 1.7	.763 ± .063	.0005182
5	(100.0)	24.7 ± 1.6	.373 ± .143	.00056
6	123. ± 22.	23.6 ± 1.4	(0.0)	.00061

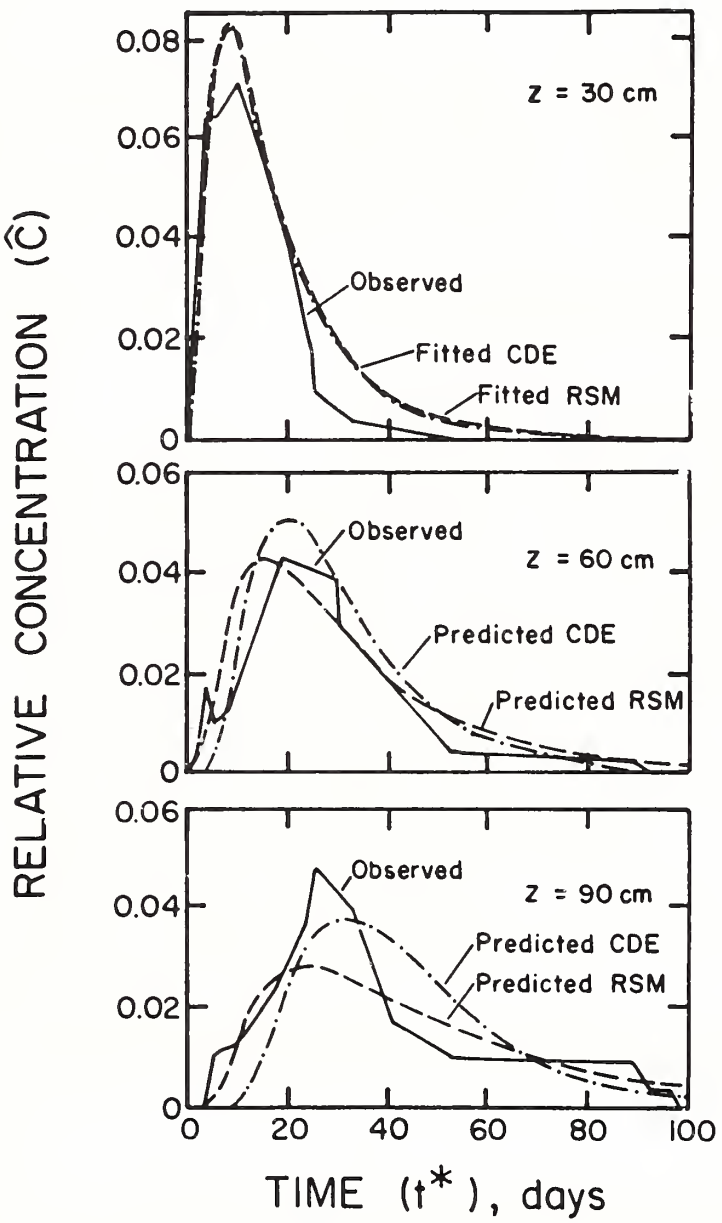


Figure 4.
Observed (Jury et al. 1982), fitted (top) and predicted (middle and bottom) areally averaged resident concentrations as a function of transformed time t^* for bromide transport in a field soil (after Kool et al. 1987).

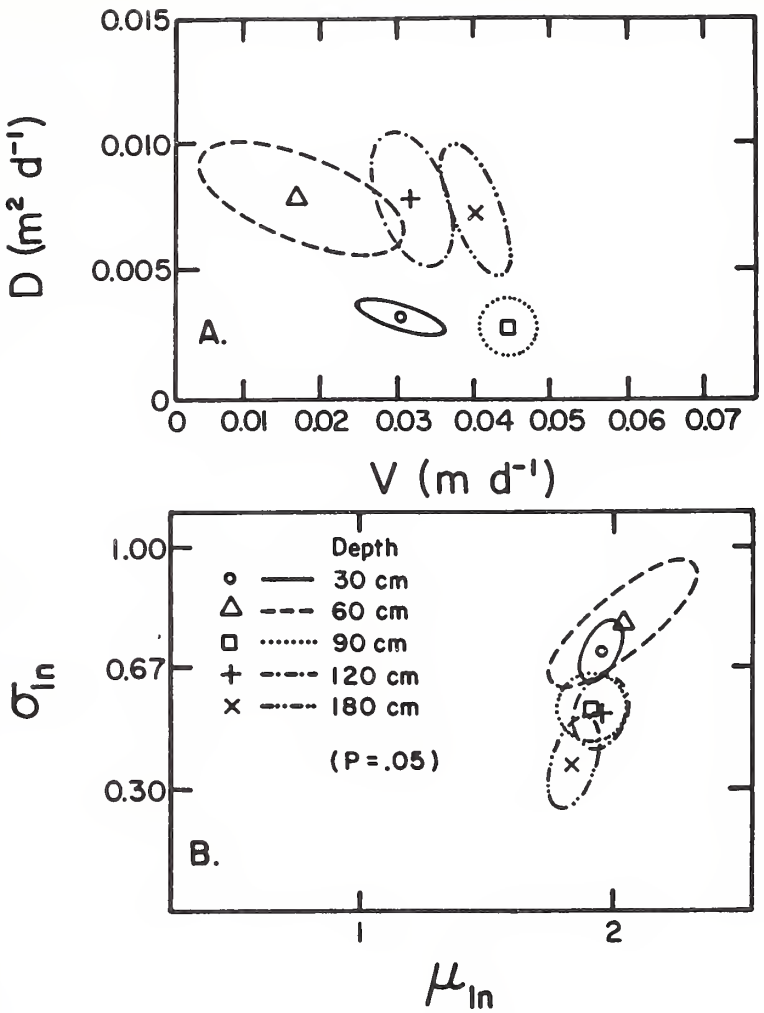


Figure 5.
Mean and 0.05 probability error ellipses for parameters fitted to breakthrough curves at different depths for the CDf model (A) and the TFM model (B) (after Jury and Sposito 1985).

$$\frac{\partial c}{\partial t} = \frac{1}{A}\frac{\partial}{\partial x}\left(AD\frac{\partial c}{\partial x}\right) - \frac{Q}{A}\frac{\partial c}{\partial x} + \frac{q_\ell}{A}(c_\ell - c) + \beta(c_s - c) \qquad [32a]$$

$$\frac{\partial c_s}{\partial t} = -\beta\frac{A}{A_s}(c_s - c) \qquad [32b]$$

where c = the river solute concentration,
Q = the volumetric flow rate,
A = the cross-sectional area of stream channel,
D = the dispersion coefficient,
c_s = the concentration of the storage zone,
A_s = the cross-sectional area of the storage zone,
β = a stream-storage exchange coefficient,
q_ℓ = the lateral volumetric inflow per unit length,
c_ℓ = the solute concentration of the lateral inflow, and
t, x = time and distance, respectively.

Note that the above dead-zone river transport model closely resembles the nonequilibrium two-region (mobile-immobile) transport models discussed earlier for transport in heterogeneous soils. Closely related applications of the dead-zone river model are given by Thackston and Schnelle (1970) and LeGrand-Marcq and Laudelout (1985), among others.

A Crank-Nicolson type finite-difference solution of equations 32a,b was coupled with a nonlinear parameter estimation procedure to identify selected mixing and storage parameters in each of five stream reaches. The parameters identified were the dispersion coefficient, the storage cross-sectional area, the average stream cross-section, and the stream- storage exchange coefficient. Four of the measured breakthrough curves from the tracer experiment, together with the model values, are shown in figure 6. The model fit is quite good. It is interesting to note that Bencala and Walters (1983) also obtained a seemingly good model fit based on visual inspection of the match between observed and simulated concentration histories. Some of their parameter values corresponded fairly well with those obtained with the simulation-regression model. However, their values for the dispersion coefficient overestimated those given by the regression model by between two and four times, showing the greatest deviations in the two downstream reaches. Furthermore, values based on manual calibration were far outside the 95% confidence intervals given by the simulation-regression procedure. We note that the use of simulation-regression models was outside the scope of the Study by Bencala and Walters. Nonetheless, this example illustrates that manual calibration can quickly lead to misleading parameter values. The value of simulation-regression and related estimation techniques lies not only in providing good estimates of the parameters, but also in giving confidence limits. Finally, the sensitivity of the parameter estimates to inclusion or exclusion of certain data can be evaluated by re-running the simulation-regression model with different subsets of the data.

CONCLUSION

As simulation models are becoming increasingly popular in research and management studies of water flow and chemical transport in saturated and unsaturated systems, the need for more

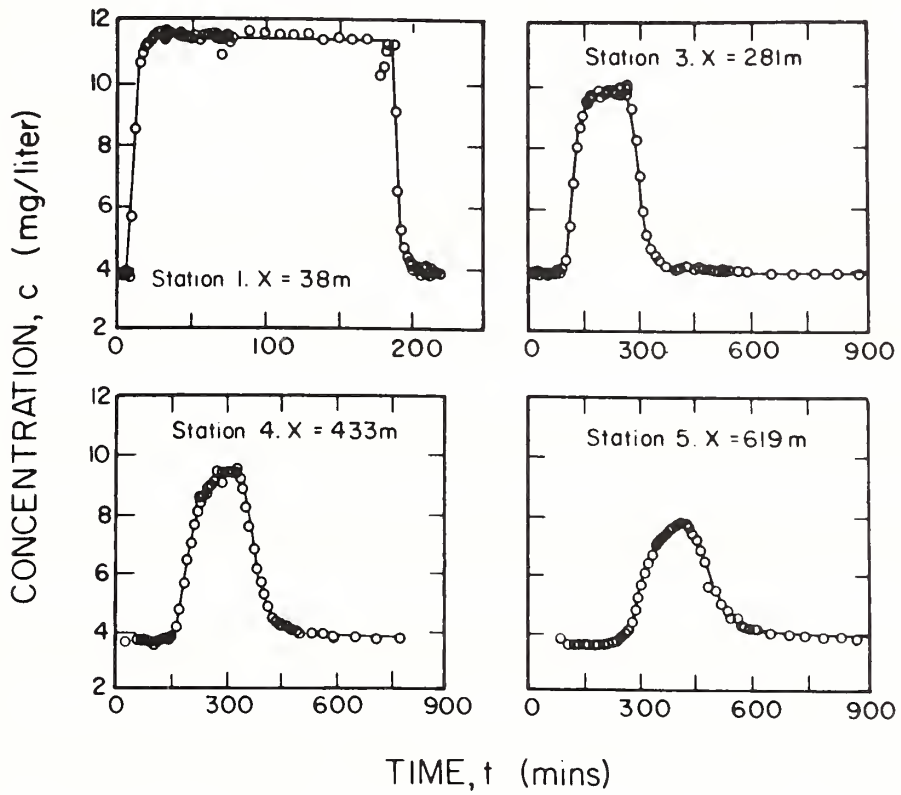


Figure 6.
Results of simulation-regression model showing measured concentrations as circles (Avanzino et al. 1984) and simulated concentration histories as solid lines at three stations in Uvas Creek. Simulation was based on best parameter values after Wagner and Gorelick 1986).

efficient and accurate methods of estimating the model parameters has become more important. Because of inherent limitations of any modeling or experimental effort, it is also important that errors associated with model parameters, and ultimately with model predictions, be quantifiable. Parameter estimation methods offer the most suitable means of meeting these requirements.

While parameter estimation methods have been widely used in saturated flow and transport modeling, they have only recently been applied to vadose zone flow and transport problems.

The majority of studies thus far have involved laboratory-scale experiments and rather simple parametric models. The few reported field-scale parameter estimation studies have involved relatively simple model formulations, boundary conditions and/or soil properties. We need to extend parameter estimation methods to more complex field conditions using models which consider such processes and properties as soil heterogeneity, variable and uncertain boundary conditions, and complex chemical and biological interactions. While studies are needed to integrate known physicochemical and microbiological processes, we also need to evaluate the point at which increased complexity and accuracy of a model is counteracted by loss of precision due to our inability to estimate a larger number of parameters with confidence. Finally, we also must come to accept that parameters are a consequence of our models; they only gain definition through solution of the inverse problem. As such, parameters are estimates, and confidence limits must always accompany their values.

LITERATURE CITED

Amoozegar-Fard A., D.R. Nielsen, and A.W. Warrick. 1982. Soil solute concentration distributions for spatially varying pore water velocities and apparent diffusion coefficients. Soil Sci. Soc. Am. J. 45:3-9.

Avanzino, R.J., G.W. Zellweger, V.C. Kennedy, and others. 1984. Results of a Solute Transport Experiment at Uvas Creek, 1972. U.S. Geological Survey, Open File Report 84-236, 86 pp.

Beck, J.V., and K.J. Arnold. 1977. Parameter Estimation in Engineering and Science. John Wiley & Sons, New York.

Bencala, K.E., and R.A. Walters. 1983. Simulation of solute transport in a mountain pool-and-riffle stream: A transient storage model. Water Resour. Res. 19:718-724.

Bresler, E., and G. Dagan. 1981. Convective and pore scale dispersive transport in unsaturated heterogeneous fields. Water Resour. Res. 17:1683-1693.

Cameron, D.A., and A. Klute. 1977. Convective-dispersive solute transport with a combined equilibrium and kinetic adsorption model. Water Resour. Res. 20:1123-1130.

Chavent, G. 1974. Identification of functional parameters in partial differential equations. In R.E. Goodson, and M. Polis, eds., Identification of Parameters in Distributed Systems, pp. 31-48, Am. Soc. Mech. Eng., New York.

Coats, K.H., and B.D. Smith. 1964. Dead-end pore volume and dispersion in porous media. Soc. Pet. Eng. J. 4:73-84.

Cooley, R.L. 1985. A comparison of several methods of solving nonlinear regression groundwater flow problems. Water Resour. Res. 21:1525-1538.

Dennis, J.E., D.M. Gay, and P.E. Welsh. 1981. An adaptive nonlinear least-squares algorithm. ACM Transactions on Mathematical Software. 7:348-368.

Gorelick, S.M., B. Evans, and I. Remson. 1983. Identifying sources of groundwater: An optimization approach. Water Resour. Res. 19:779-790.

Hiebert, K.L. 1981. An evaluation of mathematical software that solves nonlinear least squares problems. ACM Transactions on Mathematical Software. 7:1-16.

Jardine, P.M., J.C. Parker, and L.W. Zelazny. 1985. Kinetics and mechanisms of aluminum adsorption on kaolinite using a two-site nonequilibrium transport model. Soil Sci. Soc. of Am. J. 49:867-872.

Jury, W.A. 1982. Simulation of solute transport with a transfer function model. Water Resour. Res. 18:363-368.

Jury, W.A., and G. Sposito. 1985. Field calibration and validation of solute transport models for the unsaturated zone. Soil Sci. Soc. Am. J. 49:1331-1341.

Jury. W.A., L.H. Stolzy, and P. Shouse. 1982. A field test of a transfer function model for predicting solute transport. Water Resour. Res. 18:369-375.

Klute, A. ed. 1986. Methods of Soil Analysis, Part 1. Physical and mineralogical methods. Agronomy 9(1), 2nd ed. Am. Soc. Agron. Madison, Wisconsin.

Knopman, D.A., and C.I. Voss. 1987. Behavior of sensitivities in the one-dimensional advection-dispersion equation: Implications for parameter estimation and sampling design. Water Resour. Res. 23:253-272.

Knopman, D.A., and C.I. Voss. 1988. Further comments on sensitivities, parameter estimation, and sampling design in one-dimensional analysis of solute transport in porous media. Water Resour. Res. 24:225-238.

Konikow, L.F., and J.D. Bredehoeft. 1978. Computer model of Two-dimensional Solute Transport and Dispersion in Ground Water. U.S. Geological Survey Techn., Water Resources Investig., Book 7, Chapter C2.

Kool, J.B., J.C. Parker, and M. Th. van Genuchten. 1987. Parameter estimation for unsaturated flow and transport models: A review. J. Hydrol. 91:255-293.

Kreft, A., and A. Zuber. 1978. On the physical meaning of the dispersion equation and its solution for different initial and boundary conditions. Chem. Eng. Sci. 33:1471-1480.

LeGrand-Marcq, C., and H. Laudelout. 1985. Longitudinal dispersion in a forest stream. J. Hydrol. 78:317-324.

Marquardt, D.W. 1963. An algorithm for least-squares estimation of nonlinear parameters. SIAM J. App. Math. 11:431-441.

Mishra, S., and J.C. Parker. 1988. Parameter estimation for coupled unsaturated flow and transport. EOS Transactions, American Geophysical Union, 69:355.

Murty, V.V.M., and V.H. Scott. 1977. Determination of transport model parameters in groundwater aquifers. Water Resour. Res. 13:941-947.

Nielsen, D.R., and J.W. Biggar. 1961. Miscible displacement in soils: I. Experimental evaluation. Soil Soc. Am. Proc. 25:1-5.

Nielsen, D.R. , M.Th. van Genuchten, and J.W. Biggar. 1986. Water flow and solute transport processes in the unsaturated zone. Water Resour. Res. 22:89S-108S.

Nkedi-Kizza, P., J.W. Biggar, H.M. Selim, and others. 1984. On the equivalence of two conceptual models for describing ion exchange during transport through an aggregated oxisol. Water Resour. Res. 20:1123-1130.

Parker, J.C., and A.J. Valocchi. 1986. Constraints on the validity of equilibrium and kinetic transport models in structured soils. Water Resour. Res. 22:399-407.

Parker, J.C., and M.Th. van Genuchten. 1984a. Determining solute transport parameters from laboratory and field tracer experiments. Bulletin 84-3, Virginia Agricultural Experiment Station, Blacksburg, Virginia.

Parker, J.C., and M.Th. van Genuchten. 1984b. Flux-averaged and volume-averaged concentrations in continuum approaches to solute transport. Water Res. Res. 20:886-872.

Passioura, J.B. 1971. Hydrodynamic dispersion in aggregated media. Soil Sci. 11:339-344.
Selim, H.M., J.M. Davidson, and R.S. Mansell. 1976. Evaluation of a two-site adsorption-desorption model for describing solute transport in soils. In Proceedings, Summer Computer Simulation Conference, July 12-14, 1976. National Science Foundation, Washington, D. C.

Simmons, C.S. 1982. A stochastic-convective transport formulation of dispersion in one-dimensional porous media systems. Water Resour. Res. 18:1193-1214.

Sposito, G., W.A. Jury, and V.K. Gupta. 1986. Fundamental problems in the stochastic convection-dispersion model of solute transport in aquifers and field soils. Water Resour. Res. 22:89S-108S.

Sposito, G., R.E. White, P.R. Darrah, and W.A. Jury. 1986. A transfer function model of solute transport through soil. 3. The convection-dispersion equation. Water Resour. Res. 22:255-262.

Strecker, E.W., and W. Chu. 1986. Parameter identification of a groundwater contaminant transport model. Groundwater 24:56-62.

Thackston, E.L., and K.B. Schnelle. 1970. Predicting the effects of dead zones on stream mixing. Journal of the Sanitary Engineering Division, ASCE, 96:319-331.

Umari, A., R. Willis, and P. L-F. Liu. 1979. Identification of aquifer dispersivities in two-dimensional transient groundwater contaminant transport: An optimization approach. Water Resour. Res. 15:815-831.

van Genuchten, M.Th. 1980. Determining transport parameters from solute displacement experiments. Research Report 118, U. S. Salinity Laboratory, Riverside, California, 37 pp.

van Genuchten, M.Th. 1981. Non-equilibrium transport parameters from miscible displacement experiments. Research Report 119, U. S. Salinity Laboratory, Riverside, California, 88 pp.

van Genuchten, M.Th., and R.W. Cleary. 1979. Movement of solutes in soil: Computer-simulated and laboratory results. Chapter 9, Soil Chemistry, B. Physico-Chemical Models, G.H. Bolt, ed., pp. 349-386, Developments in Soil Science 5B, Elsevier, Amsterdam.

van Genuchten, M.Th., and F.N. Dalton. 1986. Models for simulating salt movement in aggregated field soils. Geoderma 38:165-183.

van Genuchten, M.Th., and J.C. Parker. 1984. Boundary conditions for displacement experiments through short laboratory soil columns. Soil Sci. Soc. Am. J. 48:703-708.

Wagner, B.J., and S.M. Gorelick. 1986. A Statistical methodology for estimating transport parameters: Theory and applications to one-dimensional advective-dispersive systems. Water Resour. Res. 22:1303-1316.

Wagner, B.J., and S.M. Gorelick. 1987. Optimal groundwater quality management under parameter uncertainty. Water Resour. Res. 23:1162-1174.

Yeh, W. W-G. 1986. Review of parameter estimation procedures in groundwater hydrology: The inverse problem. Water Resour. Res. 22:95-108.

Yeh, W. W-G., and C. Wang: 1987. Identification of aquifer dispersivities: Methods of analysis and parameter uncertainty. Water Resour. Bull. 23:569-580.

PARAMETER IDENTIFICATION

Ulrich Hornung[1]

ABSTRACT

In this paper we discuss the main aspects of parameter estimation from a critical point of view. The concepts of ill-posed and inverse problems are explained. It is shown that parameter estimation usually results as an inverse problem and that this shares the properties of ill-posed problems. Only very careful numerical treatment will in general give reliable results. The risks and uncertainties inherent in this process are pointed out.

WELL-AND- ILL-POSED PROBLEMS

Science is uncertain; the moment that you make a proposition about a region of experience that you have not directly seen then you must be uncertain. But we always must make statements about the regions that we have not seen, or the whole business is no use.

Richard Feynman, The Character of Physical Law. Cornell (1964)

Whenever a model is considered for water quality, it is in some way mathematical. This means it uses the language of mathematics to express certain relations satisfied by physical, chemical, or biological quantities. Thus, a model in most cases is an equation or a system of equations for a set of variables. The most important of these are differential equations, either ordinary or partial. Equations of this type describe the spatial and temporal behavior of the quantities in question.

First of all, talking about differential equations, there are the 'classical' problems: initial value problems (IVP) and boundary value problems (BVP) for ordinary differential equations; boundary value problems for elliptic equations, initial boundary value problems (IBVP) for parabolic and hyperbolic equations. In this kind of problems, one specifies the differential equation, all coefficients therein, and boundary and/or initial data. It has been the work of mathematicians during some 100 years to prove that many problems of this sort are well-posed. This word has a well-defined meaning: A problem is called *well-posed*, if it has all of the following three properties:

(1) existence: there is at least one solution, and
(2) uniqueness: there is at most one solution, and
(3) stability: the solution depends continuously on the data.

These three aspects are of great importance, if numerical methods are to be used. Implicitly, one takes them for granted in most cases:

(1) if there is no solution at all, it does not make much sense to calculate anything,
(2) if there are more than only one solution, it may be extremely difficult to find all of them,
(3) if the solution may change rapidly caused by slight modification of the data, any numerical result is unreliable.

Therefore, it has been the general opinion among scientists for a long time that practically all mathematical problems are well-posed, and that non-well-posed problems are irrelevant and artificial.

[1]U. Hornung, Professor, SCHI, P.O. Box 1222, D-8014 Neubiberg, West Germany.

One identifies a problem *ill-posed*, if it lacks of at least one of the above mentioned properties, i.e.

(1) non-existence: there is no solution at all, or
(2) non-uniqueness: there is more than only one solution, or
(3) non-stability: small changes of the data can cause large changes of the solution.

It turns out that there are very important ill-posed problems. One of the best-known is the diffusion equation backwards in time. Here, one studies diffusion of a chemical solute in a solvent, such as salt in water. Using conservation of mass and Fick's law, one has the parabolic equation for the concentration u

$$\partial_t u = \nabla \cdot (D \nabla u) \qquad [1]$$

Prescribing initial and boundary data for this equation leads to a classical IBVP which happens to be well-posed. But in certain applications there is information on the solution at a given time available and one wants to find out how it looked at earlier times. This problem arises, e.g., if the origin of a measured plume of a contaminant in the groundwater has to be located. In the case of the diffusion equation backwards in time, properties 1 and 3 are violated: on one hand, not any arbitrary distribution in space of the concentration u can occur; on the other hand, even the slightest change of the measured data will give rise to extremely large variations of the function at earlier times. The reason for the latter effect is that the diffusion equation forward in time has a smoothing property; even if the initial distribution has jumps, everything is smoothened out after a while; thus, going backward the problem makes smooth functions rough, which leads to an ill-posed problem.

This is the bad news. The good news is that mainly during the last 30 years applied mathematicians have designed special techniques to deal with such difficult situations. One of the main tricks being applied is now called *regularization*, a method which goes back to the Russian mathematician (Tykhonov and Arsenin 1977). Here one takes advantage of the fact that in most cases one has some *a-priori* knowledge of how the function to be determined should look, e.g., the initial distribution cannot be too oscillatory, or it must be bounded between certain quantities, etc. Using that kind of *a-priori* information, special numerical techniques have been designed to solve such ill-posed problems with sufficient accuracy. What has to be emphasized here is that if a method of textbook type that has been designed for well-posed problems is applied to an ill-posed problem, it will give nothing but garbage as pseudo-results.

Other well-known ill-posed problems are CT (Computerized Tomography), NMR (Nuclear Magnetic Resonance), and geo-electric and -acoustic seismology. In all these situations the measurements one has are obtained outside or on the boundary of some solid material, the interior structure of which has to be determined as accurately as possible. It was only during the recent years that significant progress in these fields was possible, mainly because of the very sophisticated use of hand-tailored numerical methods being run on super-computers (Natterer 1986).

THE STRUCTURE OF INVERSE PROBLEMS

Very many of the ill-posed problems that occur in practical applications are *inverse* in their structure. This means that they arise closely connected to a *direct* problem, such as a BVP or IBVP for a differential equation. To make it clear, let us consider a special example: miscible displacement of organic or inorganic materials in water flowing through an aquifer is very often described by the diffusion-dispersion-convection equation

$$\partial_t u = \nabla \cdot (D\nabla u - qu) \qquad [2]$$

Here, u is the concentration of the substance in question. If we assume that the coefficients D and q are known and given, and that initial and boundary data are prescribed, we arrive at a classical direct problem in which the function u is the unknown. This direct problem is well-posed, numerical methods have been developed that are efficient and as accurate as needed, and numerous computer codes are available for this purpose. For engineers in almost all relevant cases it is not that easy: in general they do not have all the required data needed. Therefore they are tempted to do what is often called *model calibration*. They use a computer program as a tool to play with certain parameters, such as the dispersion coefficient D. Being patient enough and having some feel for what is going on, they eventually 'match' the observed data with sufficient precision. In general they will then be satisfied with the results and use the values of the parameters found in this way for their work.

The point to be made here is this: if p denotes the parameter or set of parameters, then the direct problem will give some solution $u(p)$, in general being a function of space and/or time. Whatever kind of observations are made, these are very likely not the solution u itself, but some other quantities, say $v(u)$, such as measurements of u at certain points and at certain instances of time, or integrals thereof, also *surrogate* measurements such as stage height, which after considerable manipulation is usually reported as a measure of *runoff volume*. These, of course, are biased by noise of the measuring devices. Calling the measured data w, we are lead to the following sequence

$$p \to u \to v \to w \qquad [3]$$

This sequence describes the direct problem. The inverse problem related to it is to recover the parameter(s) p from the measurements w. Unfortunately, inverse problems of this kind have the tendency to be ill-posed in the sense of the preceding paragraph. It is very well possible that some data w do not at all belong to a possible configuration; or that different sets of parameters can yield practically the same data fit with comparable precision; or that changing w only very slightly will result in completely different parameters (Baumeister 1987). If this is so, then any model calibration has no practical significance at all.

A special example for a similar problem was discussed in the paper (Hornung and Messing 1983). That paper dealt with the determination of the *pF*-curve and the hydraulic conductivity of a soil sample in the laboratory. There the non-uniqueness aspect was discussed in detail. Later (Kool and Parker 1988; and Kool, et al. 1985) the measuring procedure was modified, such that the non-uniqueness problem was overcome. The message from this - relatively simple - example is that before one can make any statement about having solved an inverse problem, one has to go through a lengthy discussion about the direct problem. One has first to learn everything about sensitivity of parameters, both locally and globally. The latter aspect is the most tricky and difficult. It can very well be that, if one has found one set of parameters which fit the data well, say $p0$, all possible different parameters $p1$ that are small perturbations of $p0$ give worse fits. Even if this is so, it is in general not at all clear whether or not some very different parameters $p2$ may also fit with sufficient precision. Thus, the inverse problem may have a local uniqueness property without sharing a global uniqueness property. See also Dane and Bruska (1983); Carrera and Neuman (1986); and Knopman and Voss (1987).

This, in most cases, is a very complicated mathematical question for which there is no general recipe. This means that each situation has to be studied separately in all detail (Hämmerlin and Hoffmann 1983; Cannon and Hornung 1986). Going through the process of model calibration without a thorough knowledge of the difficulties involved will always lead to useless results. It should be mentioned here that, in principle, similar considerations apply in the case of stochastic

models, the only difference being that the parameters involved are statistics of random variables. A review article on parameter estimation and decision-making under uncertainty was recently written by Beck (1987), see also Kitanidis and Vomvoris (1983); Wagner and Gorelick (1986); and Lu, et al. (1988). The interested reader should have a look at the extensive literature lists provided by the papers in Beck (1987) and Yeh (1986).

THE USE OF OPTIMIZATION TECHNIQUES

From the preceding paragraph, it is clear that an inverse problem can be looked at as a special case of an optimization problem. The parameter p has to be determined such that the remainder

$$r = w - v(u(p)) \qquad [4]$$

is as small as possible (Chavent 1983). Here $u(p)$ is a short notation for the solution of the direct problem, and $v(u)$ denotes what kind of quantities are actually measured. Similarly-structured problems occur as curve-fitting problems. In this - much simpler - situation, p stands for certain parameters or coefficients of the unknown function $u(p)$ to be determined; $v(u)$ describes in most cases the evaluation of the function u at given points in time or space. The main difference to the problems of differential equations and their inverse problems is that here the direct problem is nothing but the determination of a function in terms of their parameters; this is negligible compared to solving a BVP for a nonlinear partial differential equation.

Now, once the direct problem is defined and a numerical procedure for its solution is at hand, it seems that only an optimization code has to be used in order to solve the minimization problem 4, i.e., to close the loop of 3

$$w \rightarrow p \qquad [5]$$

As simple as it looks here, as difficult it is in practice. First of all, the quantity r in 4 depends, in general, in a nonlinear way on p. Therefore, the optimization problem is not simple to be solved numerically. On the other hand, only one evaluation of r as a function of p may cost several computer hours, even on a fast machine; this stems from the fact that each time a special direct problem has to be solved, i.e., a BVP or an IBVP for a partial differential equation, which itself may be nonlinear. Assuming that computer time is not the crucial bottleneck (it becomes cheaper almost every day), one might use both modules, namely a direct problem solver and an optimization code, as 'black boxes', and plug them together, hoping that a reasonable solution will be found.

Using brute force to find a parameter fit may come with very high risks. The direct solver, in the first place, has only limited accuracy, no matter how much computer time is spent on the solution. Then, as everybody knows who has ever worked in this field, nonlinear optimization problems belong to the class of 'hard' numerical problems, in that their numerical complexity can be arbitrarily large. This is especially true if restrictions of any kind come into play, such as constraints on the state variables, etc. And, as mentioned in the last section, one has in general to deal with non-uniqueness problems. It is a well-known fact that even the simple to formulate least-squares approximation problem may have infinitely many solutions if, e.g., exponential sums of the form $u(t) = \alpha_1 e^{\lambda 1 t} ... \alpha_n e^{\lambda n t}$ are used (Braess 1986). In some cases this becomes apparent, if one starts the optimization routine with different starting points; but this need not become evident in such a simple way.

ALTERNATIVE MODELS

All kinds of parameter estimation start with a given model for the physical, chemical, or biological problem in question. But, in quite a few situations it is not at all clear which the true model should be. Very often there is more than only one competing model to be taken into consideration. This may be because the underlying process is not yet fully understood. Or because it is the decision based on which of the many possible assumptions and simplifications the modeler likes most. It cannot be emphasized too much that any model will always depend on such assumptions and simplifications. Nature is not simple. It is no doubt that there may be something like very beautiful basic laws of science. But the problems we are talking about here, namely water pollution, cannot be described in a simple way using only simple equations.

To the knowledge of the author there are few research papers which focus on the aspect of distinguishing between different type of models for the same phenomenon. The standard paper in the field of soil physics, hydrology, or water management ends with the statement that the fit of the computed numbers to the measured quantities is 'excellent'. The question of whether or not a different model would have given as 'excellent' results is often not even asked, not to mention answered. The dilemma we are in is that - in general - experimental science can never *prove* a statement, only *disprove*. Or, using the terminology of Popper (1968), experimental science can only *falsify* hypotheses. This has a very serious consequence, namely that even if the parameter estimate one has made yields very small residuals, there is no 'proof' that 'the' model has been found. It can very well be that a different approach gives as good or even better results.

One of the problems we are talking about here is the problem of scales. As soon as variables are considered that measure some quantities, certain decisions have to be or have already been made as to what their scales are. In many situations there is not 'the' scale. This is not a question of liters versus gallons or miles versus kilometers, but of non-linear transformations. As an example, let us consider the hydraulic potential in unsaturated soils. Quite a few researchers prefer a logarithmic scale for this variable instead of a linear scale. The same applies to the hydraulic conductivity. This may or may not be appropriate; we do not want to make a final statement about this here. The point is that whenever curve fitting or anything similar is applied, the results depend strongly on the underlying scales. This is also true for how to measure residuals; some people prefer maxima, others sums of squares, others use something different. There seems not to be the 'true' way of doing this.

It is not only the scales where the modeler has to make decisions; it is mainly which of the physical, chemical, and biological mechanisms that may influence the processes to be modeled are taken into account and which are neglected. Unfortunately, if an effect is - *per se* - small, it may nevertheless at the end very well have dramatic consequences. Let us only mention here that pollutants - even in extremely small concentrations - may have disastrous implications. And if their temporal kinetics is underestimated only by some small percentage, the discrepancy to reality can become enormous. Therefore, one has to be aware from the very beginning of the modeling process that basic decisions are being made which may make the whole work irrelevant.

This continues through the whole work, namely at each stage there are very often alternatives to be considered. One of the most frequently ignored difficulties is that of how to formulate boundary conditions. Modeling a process that takes place in space makes it always necessary to cut a fraction out of the universe, namely the domain of interest. Now, there is no means to do any simulation without specifying the connection of this domain to the rest of the world. But, in many cases this is not simple. Often enough, there is practically no knowledge about these phenomena. On the other hand, the solution of a BVP or IBVP strongly depends on the boundary conditions chosen. Thus, different choices made at this point may yield different results at the end.

Another crucial point is that of linearization. Evidently, linear differential equations are much easier to handle than nonlinear ones and much more is known about them, such as solutions in form of power or Fourier series, etc. And, obviously, linear problems allow numerical approximations with much less effort than nonlinear problems. But, first of all, nature is nonlinear. Secondly, the linearization may have different mathematical properties than the nonlinear version that it is derived from. This applies under any aspect, such as well- or ill-posedness, well- or ill-conditionedness, or any kind of stability or instability. Without actually performing a careful analysis, both theoretically and numerically, nothing can be said about the possible consequence of replacing a nonlinear problem by its linear counterpart. Very often every important feature of the original problem may be lost. Unfortunately, many textbooks on applied mathematics do not even mention this very basic aspect.

It would make research papers much easier to read for the engineer who wants to use the results, if the underlying assumptions and simplifications were stated in a precise and easy to understand style.

MORE RECENT MODELS

Soil physics, chemistry and biology are relatively young scientific disciplines. Essentially, these branches of science are not older than some one hundred years. Therefore, it cannot be a surprise that we are in a rapid increase of knowledge and methodology. What was the latest news yesterday may be obsolete today; and what is fashionable today may become irrelevant tomorrow. Looking into textbooks on soil science and related fields and checking what kind of mathematical tools are being used, one finds out very quickly that there is a significant time lag between what is now used in the engineering sciences and what modern applied mathematics has to offer. Here we mention only some of the more recent concepts of research in soil science.

Unconfined aquifers give rise to *free or moving-boundary problems* (Hornung 1989). This means that the phreatic surface is the boundary of a domain, the shape of which has to be determined as part of the problem. Within the domain filled by water there is a partial differential equation given; on the phreatic surface there are two conditions, namely one on the pressure and one on the normal flux. It is only the fact that there are these two pieces of information which makes the whole problem well-posed. The theory of this type of problems is now well-developed without having much effect on practical applications. Most simulations are still based on something like the 'Dupuit assumption', an assumption everybody knows is incorrect (Hornung and Krüger 1985).

One of the commonplaces in soil science nowadays is that there are *fractured media*, i.e. media which have two types of conductivities - or permeabilities - simultaneously. It turns out that the usual approach, namely the standard Darcy law, does not describe this kind of media. The type of differential equation resulting from Darcy's law is not appropriate to model fractured or fissured media. There are new models available that take into account the effect of double porosity (Arbogast, et al. 1989; Arbogast, et al. 1988; Hornung 1989 and; Hornung and Showalter 1989).

Another important aspect of modern research is that of *spatial variability*. Measuring any quantity in the field gives large variations of the variables; and this is - as the experimentalists seem to agree - independent of the length scale. There has been quite some effort to use all methods from descriptive statistics to evaluate these observations. But very little has been done so far to combine this intelligence with what is known about the physical and chemical processes and their laws. Thus, one of the big open questions is that of solving *random partial differential equations*, i.e. equations with coefficients that are random fields (Dikow and Hornung 1988).

It should be pointed out very clearly that in case there are two *competing models* which give comparable results, the one to be preferred is the simpler, which uses the more elegant tools, etc. But here the premise is that of comparable results. Unless one has actually checked all possibilities and gone through the whole analysis, nothing can be said about the value or usefulness of any assumption and method. The aspects should never be that of taste or subjective belief, but only that of performance and efficiency measured using objective criteria.

As long as the above mentioned more recent aspects and results do not find their way to applications, modern science has not been used at its full power. And as long as these concepts do not compete with more traditional approaches, the full variety of models is not taken into serious consideration. Even the very best parameter estimation technique is of little use if the underlying model is poor (Hornung 1986). And, finally, there are no "general purpose" models.

OPEN PROBLEMS

This may be also the place to mention several aspects of flow and transport through porous media for which basic questions are still open. As long as it is not clear what a reasonable model for a certain process is, it remains doubtful whether the problem of parameter identification makes much sense. There are quite a few mechanisms which scientists do not yet fully understand. Among these are the following (naturally, this list is very subjective, influenced by the author's prejudices).

First of all, for two- or multiple-phase flow there is no justification for the concept of *relative permeabilities*, as used by many people. It seems not to be clear how to derive Darcy's law from more basic principles in this case. Also, the relation between the capillary pressure and the saturation has not been put on a sound basis. In this context, the effect of *hysteresis* is not completely understood.

There are many attempts to model fractured media, as it was pointed out in the preceding chapter. Different from this is the phenomenon of *macropore flow* - a situation in which the flow through the large pores does not obey Darcy's law. This is of great importance in the event of heavy storms, and it significantly influences the transport of solutes. Similarly, even though detailed descriptions of experimental studies have been published on the problem of *fingering* and *preferential flow*, this process has not yet been modeled adequately.

Hydrodynamic dispersion and *macro-dispersion* are well-established facts of experimental soil science. Up to now there does not seem to be an answer to the question of how to explain the mechanisms which cause these effects. And it is still an open question to quantify the dispersion coefficient and its dependence on the Darcy velocity.

The physical, chemical, and biological processes that influence *evapotranspiration* are extremely complex. Therefore, scientists are far from having a reliable model to describe what takes place between the surface of the earth, the plants, and the atmosphere. Even the simple question of how to measure the hydraulic conductivity in the *root zone* does not have an answer. Similarly, in the situation of *simultaneous open channel and subsurface flow* it is an open problem to define the interface conditions between the two flow regimes.

Finally, as soon as chemical reaction kinetics and the dynamics of biological phenomena come into play, there are hardly any well-established facts. It is especially difficult to determine *sorption* and *reaction rates*, both experimentally and theoretically. What is known from classical chemistry about the interaction of various species cannot directly be applied to porous media. One of the

fundamental difficulties is that some of the toxic pollutants have extremely slow reaction and *degradation rates*, a fact which makes it almost hopeless to measure these quantities accurately.

The mathematical theory of soil physics and soil chemistry turns out to be a challenge. Not only complex processes are to be modeled but also such questions as different *scales* - both in space and in time - are of importance. What takes place within pores of the size of fractions of millimeters has an effect on what one observes on the scale of kilometers. The geometry of porous media seems to have the nature of fractals. *Heterogeneities*, properties that are essentially stochastic, large systems in which hundreds of substances influence each other are the basic obstacles on the way to understanding and modeling soils and what is going on in seepage and groundwater.

CONCLUSIONS

There exist quite a few numerical techniques for determining parameters in a given mathematical model of soil science. The main features of the procedure of parameter estimation are that

(1) the problem - for a given model - is ill-posed,
(2) the solution depends strongly on the type of model chosen.

The first of these properties makes the use of special numerical methods mandatory; these are available today. The second property causes more difficulties; more basic research is needed to understand the dependence of predictions upon the model assumptions. One cannot substitute lack of theory and/or data by sophisticated mathematical methods for parameter identification.

If you've made up your mind to test a theory, or you want to explain some idea, you should always decide to publish it whichever way it comes out. If we only publish results of a certain kind, we can make the argument look good. We must publish both kinds of results.[1]

ACKNOWLEDGEMENT

This work was partially supported by the United States Department of Agriculture.

REFERENCES

Arbogast, T., Douglas, Jr., J. and Hornung, U. 1989. Derivation of the double porosity model of single phase flow via homogenization theory. SIAM J. Mathem. Anal. Submitted.

Arbogast, T., Douglas, Jr., J. and Hornung, U. 1988. Modeling of naturally fractured petroleum reservoirs by formal homogenization techniques. Technical Report Center for Applied Mathematics, Purdue University 81.

Baumeister, J. 1987. Stable Solutions of Inverse Problems. Vieweg, Braunschweig.

Beck, M. B. 1987. Water quality modeling: A review of the analysis of uncertainty. Water Resources Research 23: 1393-1442.

Braess, D. 1986. Nonlinear Approximation Theory. Springer, New York.

[1]Richard Feynman, Commencement Address, CALTECH (1974).

Cannon, J. and Hornung, U. 1986. Inverse Problems. International Series of Numerical Mathematics 77. Birkhäuser. Basel.

Carrera, J. and Neuman, S.P. 1986. Estimation of aquifer parameters under transient and steady state conditions 3: Applications to synthetic and field data. Water Resources Research 22: 228-242.

Chavent, G. 1983. Stability of parameter estimates by the output least square method. H. Hämmerlin, K.-H. Hoffmann (Eds.) Improperly Posed Problems and Their Numerical Treatment. Birkhäuser. Basel. 37-45.

Dane, J.H. and Hruska, S. 1983. In-situ determination of soil hydraulic properties during drainage. Soil Science Society of America Journal 47: 619-624.

Dikow, E. and Hornung, U. 1988. A random boundary value problem modeling spatial variability in porous media flow. M.F. Wheeler (Ed.) Numerical Simulation in Oil Recovery. The IMA Volumes in Mathematics and Its Applications 11. Springer. New York. 111-117.

Hämmerlin, H. and Hoffmann, K.-H. 1983. Improperly Posed Problems and Their Numerical Treatment. Birkhäuser. Basel.

Hornung, U.: 1986. The interface. On the interaction between applied mathematics and the engineering sciences. Universität der Bundeswehr München, Fakultät für Informatik, Bericht 8610.

Hornung, U. 1989. Compartment and micro-structure models for miscible displacement in porous media. Rocky Mountain J. Math. In press.

Hornung, U. 1989. Free and moving boundary problems arising in porous media flow. K.-H. Hoffmann (Ed.) Free Boundary Problems. Theory and Applications 5 Pitman. In press.

Hornung, U. and Krüger, Th. 1985. Improved formulas for the shape of the phreatic surface with accretion. Water Resources Research 21: 1494-1496.

Hornung, U. and Messing, W. 1983. Identification of nonlinear soil physical parameters from an input-output experiment. P. Deuflhard, E. Hairer (Eds.) Numerical Treatment of Inverse Problems for Differential and Integral Equations. Birkhäuser. Boston. 227-237.

Hornung, U. and Showalter, R. 1989. Diffusion models for fractured media. J. of Mathem. Anal. Applics. In press.

Kitanidis, P.K. and Vomvoris, E.G. 1983. A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulation. Water Resources Research 19: 677-690.

Knopman, D.S. and Voss, C.I. 1987. Behavior of sensitivities in the one-dimensional advection-dispersion equation: Implications for parameter estimation and sampling design. Water Resources Research 23: 253-272.

Kool, J.B. and Parker, J.C. 1988. Analysis of the inverse problem for transient unsaturated flow. Water Resources Research 24: 817-830.

Kool, J.B., Parker, J.C. and van Genuchten, M.Th. 1985. Determining soil hydraulic properties from one-step outflow experiments by parameter estimation. I. Theory and numerical studies. Soil Science Society of America Journal 49: 1348-1354.

Lu, A.H., Schmittroth, F. and Yeh, W.W-G. 1988. Sequential estimation of aquifer parameters. Water Resources Research 24: 670-682.

Natterer, F. 1986. The Mathematics of Computerized Tomography. Teubner, Stuttgart.

Popper, K. 1968. The Logic of Scientific Discovery. Harper and Row. New York.

Tykhonov, A.N. and Arsenin, V.Y. 1977. Solution of Ill-Posed Problems. Halsted, Washington.

Wagner, B.J. and Gorelick, S.M. 1986. A statistical methodology for estimating transport parameters: Theory and application to one-dimensional advective-dispersive systems. Water Resources Research 22: 1303-1315.

Yeh, W.W.-G. 1986. Review of parameter identification procedures in groundwater hydrology: The inverse problem. Water Resources Research 22: 95-108.

DISCUSSION OF THE PAPERS PRESENTED IN TECHNICAL SESSION 8, PART 2: PARAMETER IDENTIFICATION

Daniel Hoggan[1], Presiding
Donna Falkenborg[2], Recorder

PAPERS DISCUSSED

Application of Parameter Estimation Techniques to Solute Transport Studies by M.Th. van Genuchten, S.M. Gorelick and W.W-G. Yeh

Parameter Identification by U. Hornung

SPECIFIC QUESTIONS AND COMMENTS

Question: (I. Baker, USDA-ARS, St. Paul, Minnesota) The criteria for a well-posed problem, you mentioned, all seemed to be defined in terms of the solution. They offer no help in deciding if you have a well-posed problem. Can you offer some help?

Response: (U. Hornung, Department of Mathematics, Arizona State University, Tempe, Arizona) There are no general guidelines. That's the question you have to deal with in all detail in a given situation and it may take years to research.

Question: (J. Baker) In multiple-parameter estimation procedures, how do you satisfy yourself with regard to the uniqueness of your solution when you do get an answer?

Response: (M. van Genuchten, USDA-ARS) First, I would run the estimation process with different parameters and with different initial estimates to see if they converged to the same global minimum, and then use methods to derive independent estimates of those parameters; and use judgement whether or not any of those things make sense physically.

Question: (D. Jackson, Susquehanna River Basin Commission, Harrisburg, Pennsylvania) In ordinary analysis of least squares regression, the techniques for error analysis are defined and documented. Are those techniques similarly well-defined and similarly applicable in the kinds of identification problems you talked about?

Response: (U. Hornung) There is a pretty well-developed theory now about linear inverse problems that has error analysis. One is able to combine the aspects of dispersion errors of the direct problem on one hand with the errors which remain as noise on the other. You have to find some kind of balance. Little has yet been done to carry this insight over to nonlinear problems.

Question: (D. Woolhiser, USDA-ARS, Tucson, Arizona) Would you comment on using some preliminary analysis on the finite difference methods as an aid in designing experiments that will use the data for parameter estimation?

---

[1]Daniel Hoggan, Professor, Utah Water Research Laboratory, Utah State University, Logan Utah.
[2]Donna Falkenborg, Logan, Utah.

Response: (U. Hornung) It is good advice before performing your experiment in a numerical simulation run, to have a classroom game in testing the whole routine. You define certain measurements at certain points, go through the whole analysis assuming some bias data and apply your numerical technique, then go through your sensitivity analysis and try to find out how reliable your results are. Play the same game with different measurement/observation rules. It would save time if more experimentalists would do this before they go to the lab or field and take measurements.

Comment: (M. van Genuchten) There is a lot of work to be done in this area, but basically I look at the derivatives with respect to certain parameters of concentration, and it gives hints of where to sample and how dense to sample. This is a major challenge not to focus as much on the statistics of existence and stability problems, but to look at the type of boundary conditions and initial conditions that give us the highest sensitivity. So it's an experimental challenge as well as a statistical challenge.

Comment: (D. Woolhiser) I found this to be very true in terms of estimating parameters for erosion equations by field simulation techniques. By going through some of these exercises we decided to change some of the techniques that would have been proposed because the parameters simply would not have been identifiable.

Comment: (U. Hornung) It was a couple of years ago in Germany when we did the same kind of theoretical study for very simple dynamic measurement of hydraulic conductivity of supposedly homogeneous soil. We found that in a certain lab experiment structure you would have very strong non-uniqueness. We made these simulations showing that practically the same outflow could be modeled with nearly the same convincing accuracy, for the experimentalist, using different choices of parameters. Then in a following paper changed the experimental setup and showed that under different conditions this non-uniqueness disappeared. This was a simple kind of textbook example. The principle applies to the more complicated problems.

Comment: (D. Aum, Missouri) In groundwater modeling we are more advanced and are doing a little better than what we have seen. In parameter identification estimation, by using statistical methodology like maximum likelihood we get the mean and variance of estimation of parameters. If you have an objective function like least squares you examine not only the values which give you the minimum but also the derivative and the secondary derivative function near the minimum. If the surface response is very flat, it will show you that your parameters are subject to a large variance. If you have more parameters you can make a best fit, but this doesn't mean that you have gotten the best values or have the best numbers. There are techniques which have been applied in parameter estimation which give more than just a best fit of a figure, with measured observations. The second topic is much more serious. It has been pursued in groundwater flow for 15 years, and in the last few years we have become more convinced that a better way to set the problem is in the stochastic framework. It is foolish if you have noisy data to try to find parameters so the model will perfectly fit the data. So you say what I have is a stochastic random variable and what I want to find is only its mean and some statistical moments, which can be determined in a much more stable way than the parameters themselves. Work has been done in this area and I think it is quite promising.

Response: (U. Hornung and M. van Genuchten) We have also used the maximum likelihood method. Of course we get the ground estimates and the confidence intervals of the uncertainty associated with them. We have also used some statistical formulations; but they are quite simple approaches and not of the degree of sophistication that you have worked on.

Comment: (T. Jakeman, Australian National University, Australia) There aren't a lot of differences between what you are both saying. A traditional mathematical technique for handling these problems is regularization and it sort of looks deterministic because you inject some

information to stabilize the least squares formulation. It tends to take the form of derivative-smoothing on the solution. You tend to trade-off smoothness in solution resolution against fit of the data and try to find a balance. Now the thing with techniques that have been used in groundwater-flow modeling, is that one can inject probability distributions as prior information about the form of the parameters and there is a one-to-one correspondence between both formulations. They are doing exactly the same thing. They just have different interpretations.

Questions: (D. Hoggan, Utah Water Research Laboratory, Utah State University, Logan, Utah) I think the point you made regarding the selection of the model and its importance has a couple of facets. I know in the area of hydrology and hydraulic modeling there are a number of models available but it's an investment in time to try to understand these other models. Do you have any suggestions as to: 1) how do you identify other models in a reasonable practical way that might be more applicable to a particular solution, and 2) how do you overcome the inertia of staying with what you are used to?

Response: (U. Hornung) I want to point out very clearly that you have a contrast. On one hand, once you have selected a model then solution would be only a mathematical problem. I can do nothing more than emphasize this aspect as much as possible. Choosing different assumptions from the very beginning, when you set up the model, or choosing different simplifications, you always end up with different kinds of mathematical problems which may have very different properties and the outcome at the end will depend strongly on these assumptions. No matter how much effort you put on optimizing parameters in the mathematical framework, if you made a big mistake in the beginning you can never overcome that. I don't know how to deal with the reluctance to deal with unknown things. I have no recipe for that.

Question: (C. Duffy, Department of Agronomy, Cornell University, Ithaca, New York) Can a mathematician live with the role of physical intuition in solving problems?

Response: (C. Hornung) A fine mathematician lives in some kind of conflict. Mathematics itself is based on assumptions; I use rules of logic to deduce other statements, prove theorems, and so forth. Everything related to natural science where experiments are involved is of a completely different nature. Assumptions are no longer, as in mathematics, arbitrary. A mathematician can switch from this assumption to the next, where a physicist or chemist cannot.

Response: (M. van Genuchten) You must recognize that any model is a tool for doing work and you're the one that is going to use the model. You can never decipher away your judgment, your expertise.

Comment: (U. Hornung) It might be a nice idea to set up some kind of competition. Choose a well-defined situation, say, on solid transport. Provide experimental data, then ask a group of modelers to estimate parameters and to make a prognosis about what this special setup will give under different conditions, and I think it would be a big disappointment to many participants to find out their prognosis is not found in reality. That was my last point. In a mathematical sense you can in some way say this is my optimal fit, but this is only in the framework of mathematics. I think the user and the decision-maker have completely different problems, and I think the question is, "What is a good model for a certain situation?" That is not a mathematical question. Maybe this kind of competition could somehow clarify the problem.

Comment: (C. Duffy) We are involved in such a case in Los Alamos. They carried out some large caisson experiments and then invited about 6 or 7 researchers from around the world to come and evaluate the data and have a so-called bloodbath. It didn't turn out that way. Maybe the conditions were such that it was too easy to analyze the data but we all came out with

reasonably similar results. A simple caisson experiment of tracers moving through a few meters of soil. Maybe there is some confidence in that.

Comment: (B. Kelly, University of Nebraska, Lincoln, Nebraska) There are a lot of these competitions in Superfund. It's too bad that it won't get in the literature because of the legal implications. A comment on inverse problems: in many areas it's much more highly-developed, geophysics for example. Remote sensing in geophysics problems, by their very nature, are only interested in the inverse problem and there are very simple rules of application that can be carried over into groundwater hydrology applications.

USE OF SOIL SURVEY AND OTHER DATA BASES IN THE MODELING OF LEACHING FROM AGRICULTURAL SOURCES

A. Breeuwsma[1] and J. Bouma[2]

ABSTRACT

Application of models that predict the regional effects of pollutants being leached from the root zone in agricultural soil requires data bases on climate, soils, land use and chemical application. The paper describes development and use of these data bases. Particular emphasis is given to the use of soil survey data in deriving model parameters (by so-called pedo-transfer functions) and spatially representative values. Three case studies dealing with the leaching of nitrate and phosphate illustrate the approach and the application of digitalized data bases derived from soil maps and remote sensing.

INTRODUCTION

Water quality modeling is an essential tool in predicting the effect of protection policies and management alternatives on leaching from agricultural soils. To date, research in this field primarily addresses model development and improvement. Comparatively little attention is being paid to the development and use of data bases needed for the application of models on a regional scale. Adequate data bases and methods for collecting representative data are equally important as model formulation in obtaining reliable predictions. We need data bases on, for example, climate, soils, land use and such agricultural practices as fertilizer and manure application. The soil data base is of special importance because soil properties have a high spatial variability and cannot easily be altered. Land use also varies in space but can be changed to a certain extent. Climate is less variable and the use of fertilizers, manures, pesticides, etc. is strongly affected by changes in agricultural management and environmental protection strategies.

Therefore, in describing the state of the art in data base development and use, particular emphasis is given to:

(1) the derivation of model parameters from soil survey data
(2) the use of sampling methods, soil maps and/or geostatistical techniques to ascertain the spatial distribution of model parameters
(3) recent case studies on the use of (digitalized) data bases in modeling the leaching of nitrate and phosphate on a regional scale.

DATA BASES FOR MODEL DEVELOPMENT AND USE

Data required for the development and use of leaching models may be classified as input data (driving variables), state variables, and parameters (table 1). Input data involve the inputs of water and solutes. Parameters are constant coefficients or relations in mathematical functions. They can usually be related to physical characteristics of the system that have a constant value or distribution within a temporal and spatial unit, like soil or crop descriptors.

[1]A. Breeuwsma, Head of the Soil Chemistry Division, Netherlands Soil Survey Institute (STIBOKA), Wageningen, The Netherlands

[2]J. Bouma, Professor of Soil Inventory and Land Evaluation, Agricultural University, Wageningen, The Netherlands

Table 1.
Examples of data and geographic data bases used in leaching models for agricultural soils.

Data	Examples	Databases
Input data		
- weather	water input	meteorological
- agricultural chemicals	solutes input, e.g., fertilizers, pesticides, and manure	agricultural statistics
- crop/land use	precipitation surplus use maps	remote sensing, land
State variables	solute concentrations	---
Parameters		
- soil	water retention curves	soil survey
	hydraulic conductivity curves	ditto
	adsorption characteristics	ditto
	rate constants	ditto
- crop	rate function for uptake of nutrients (varies with maturity)	---

Model Development

Model development, including calibration and validation, needs information on all three types of data. Data requirements vary substantially with model type, complexity, and structure. Therefore, data sets collected for one model are often not appropriate for calibration or validation of another model. Thus standardization of monitoring procedures and data sets, as envisaged in a Joint European Research Project of the EEC on Nitrate in Soils is a must for (international) cooperation in model development. In addition, data sets for model development should be systematically stored in (inter)national data bases to further enhance the testing of models in different situations. The advantage of a more systematic monitoring of field data on sites with representative climate zones, land use and soils has been recognized by many, including the Soil Survey of England and Wales (Robson et al. 1987). A monitoring program has been set up to study the leaching of nitrate from agricultural soils. The quantity of data measured is not yet sufficient to test detailed process-oriented models but the monitoring program may be further extended.

Model Use

Use of (validated) models relies primarily on data bases for input data and parameters and, to a lesser extent on data bases for variables. The amount of data required varies from low for field models with lumped parameters to very high for regional models with spatially-distributed parameters and input data. Depending on the application, the geographic information should include data on weather, land use, soils, fertilizers, manures, pesticides, etc. (table 1). *Meteorological* data bases are usually available on a national scale. Many countries also have, at least for some areas, (computerized) data bases for *soils* based on mapping scales of 1:50 000 to 1:250 000. For example, in the Netherlands the Soil Information System (BIS) has both spatially identified point data and areal data which are based on representative profiles defined by the soil surveyor (Bregt et al. 1986). For the EEC a data base is being developed on the scale 1:1,000,000. Data bases for *land use* may be derived from topographical maps. However, this does not provide information on crops, and maps quickly become outdated because of changes in land use after mapping had been done. To date, actual information on crops and natural vegetation may be readily obtained from remote sensing images following calibration against the "ground truth" (Van der Laan et al. 1987). This approach is now used in the Netherlands in the mapping of risk areas for phosphate leaching (Breeuwsma and Schoumans 1987). The geographic distribution of rates of application of fertilizers, manures and pesticides usually has to be assessed from production figures and farming practices. Quantities *of fertilizers* are probably the least difficult to estimate because farmers tend to confine the rates for economic reasons. The rates may thus be estimated by assuming "good farming practices". *Manures*, however, may be applied at much higher rates than are required for optimal crop production. The application rates of manure varies widely from farm to farm and from field to field. In the Netherlands, over use can be assessed to some extent because animal stocking rates are recorded every year. At present, data are available for agricultural areas of 10,000 - 100,000 ha and in the near future for regions of, hopefully, about 400 ha (2 km^2). With *pesticides* the situation is very difficult because turnover figures are not available on a regional scale.

The combination of models with a Geographic Information System (GIS) and a Data Base Management System (DBMS) provides a modern base for a computerized groundwater vulnerability assessment system, as shown in figure 1. This is further enhanced by the rapid increase in computer capabilities which no longer prohibit the use of models with spatially-distributed parameters. Some advantages of these models over lumped-parameter models have been described by Abbott et al. (1986). Spatially-distributed parameter models divide an area into spatially homogeneous units and allow the use of physically-based parameters, less extensive calibration, and identification of vulnerable areas and their contribution to the total response of the region. To apply these models on a field or regional scale a GIS should contain data that can be related to model parameters and represent spatially homogeneous units. Stochastic models do not need spatial data but require information on the distribution of parameter values within an area and, therefore, data bases.

USE OF SOIL SURVEY DATA IN PARAMETER DERIVATION

In most cases, soil survey data listed in a GIS cannot be used directly in leaching models because they are not used as model parameters. For example, organic matter and texture are routinely measured or estimated in the field but more complex data used in models including hydraulic conductivity curves and adsorption data (selectivity coefficients and sorption rates and capacities) are much less documented. It is essential, there-fore, to ensure that relations exist between easily measurable soil characteristics stored in GIS and soil properties that are more difficult to obtain. Examples of the approach developed by the Netherlands Soil Survey Institute have been described by Breeuwsma et al. (1986) and Bouma et al. (1986). The relations between soil (land)

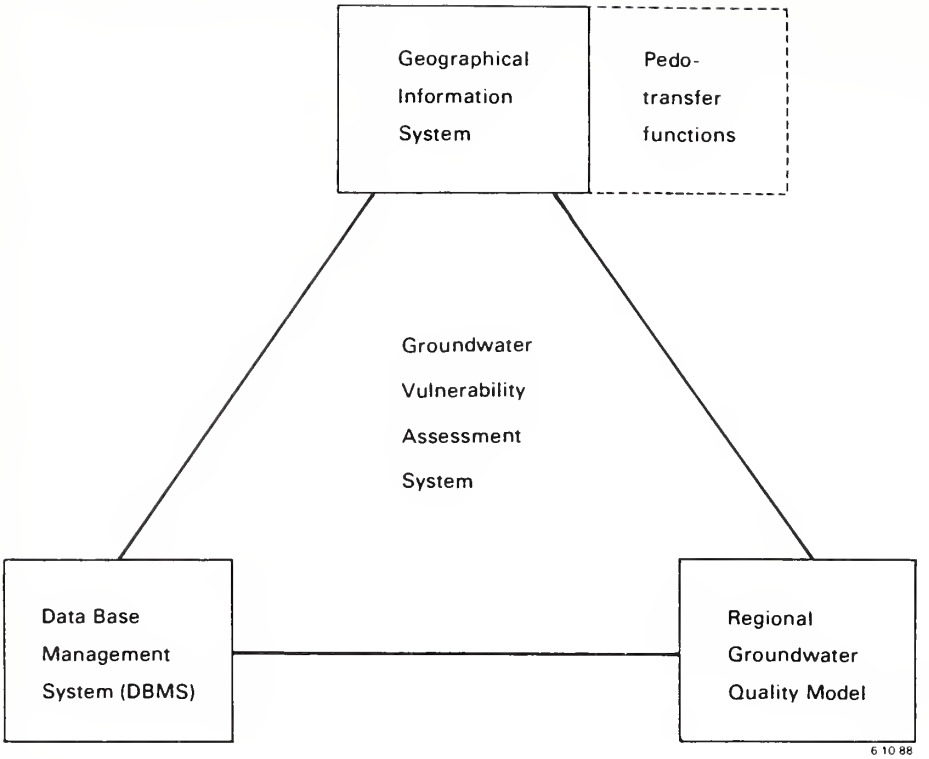


Figure 1.
Conceptual diagram of a computerized groundwater vulnerability assessment system.

characteristics and soil properties or land qualities used as model parameters were hitherto referred to as "transfer functions" (Bouma and van Lanen, 1987). However, this term is also used with a different meaning (e.g., Jury 1982) and to avoid misunderstanding we will refer to these relations as "*pedo-transfer functions*", modified after Lamp and Kneib (1981). The terminology used is described in table 2.

The pedo-transfer functions (PTF) play a central role in the leaching models developed in The Netherlands (last section and de Vries et al, in press). One of these models was set up to simulate the phosphate transport in heavily manured soils (Breeuwsma and Schoumans 1987). Figure 2 illustrates for this application how the relevant model parameter, the total phosphate sorption capacity (PSC) of the unsaturated zone, can be derived from soil survey data. The continuous function (PTF 1) relates the total PSC per unit mass to oxalate extractable Al and Fe (Schoumans et al. 1987). The assessment of the PSC is not simple, in contrast to the measurement of Al and Fe. The latter soil characteristics are therefore routinely measured during soil surveys.

Their values tend to vary with soil series (legend unit and geological formation) and horizon designation as a result of the pedogenic background of the Dutch soil classification system. The PSC can, therefore, also be derived by class functions (PTF 2 and 4). Obviously, this requires sufficient accuracy of the *soil map* with respect to the model parameter and a proper definition of the profiles representing a mapping unit. Research is underway to test this particular application of soil maps using geostatistical techniques as described for the collection of physical data (next section).

Table 2.
Terminology used in relating soil survey data to model parameters (modified after Bouma and van Lanen 1987).

Data	Definition
soil characteristics	static attributes of pedons providing information on soil composition, structure and topography
land characteristics	static attributes of areas of land
continuous characteristics	attributes that have a continuous range of values (e.g. % clay)
class characteristics	attributes characterized by a (narrow) range of values (e.g. texture class) or by a symbol (e.g. horizon designation)
soil/land properties	static attributes that can be related to soil/land characteristics (e.g. water retention curve, cation exchange capacity)
land qualities	dynamic attributes that can be related to soil and land characteristics and properties (e.g. water-table level, soil water deficit)
pedo-transfer functions	relations between model parameters and -continuous characteristics: continuous PTF -class characteristics: class PTF

EFFECT OF SPATIAL VARIABILITY ON COLLECTING REPRESENTATIVE SOIL DATA

Sampling Methods

Collecting representative soil data for areas of land presents the problems of obtaining representative point data at specific locations and extrapolating to well delineated areas of land. The latter should be relatively homogeneous internally with regard to the particular data being considered, while differences with other adjacent areas should be significant. The topic of spatial variability is a popular one within soil science, particularly in relation to physical data which are strongly affected by soil structure. However, emphasis is usually placed on (geo)statistical manipulation of data, rather than on procedures by which data are obtained. Both aspects will be covered in the following discussion, with special emphasis on physical data. Spatial variability is being recognized on the basis of variation among multiple measurements within a particular area of land, using a particular method. Measurements may involve removal of samples from a soil profile to be analyzed in the laboratory, or placement of *in situ* equipment at a selected location in the landscape and within the soil. Sampling therefore involves different aspects that cover the selection of: (1) method; (2) sample dimensions where applicable; (3) sampling locations inside the soil profile; (4) number of replicates; (5) time of sampling; (6) irregular flow patterns and (7) sampling locations in the field. Proper consideration of these aspects will result in more representative data. Often, too little attention is paid to these sampling aspects and data are collected without due consideration of method of selection and soil conditions. Data thus obtained have high variability that has only a remote relationship to the soil property being characterized. The main purpose of using proper sampling techniques is to obtain representative data for a

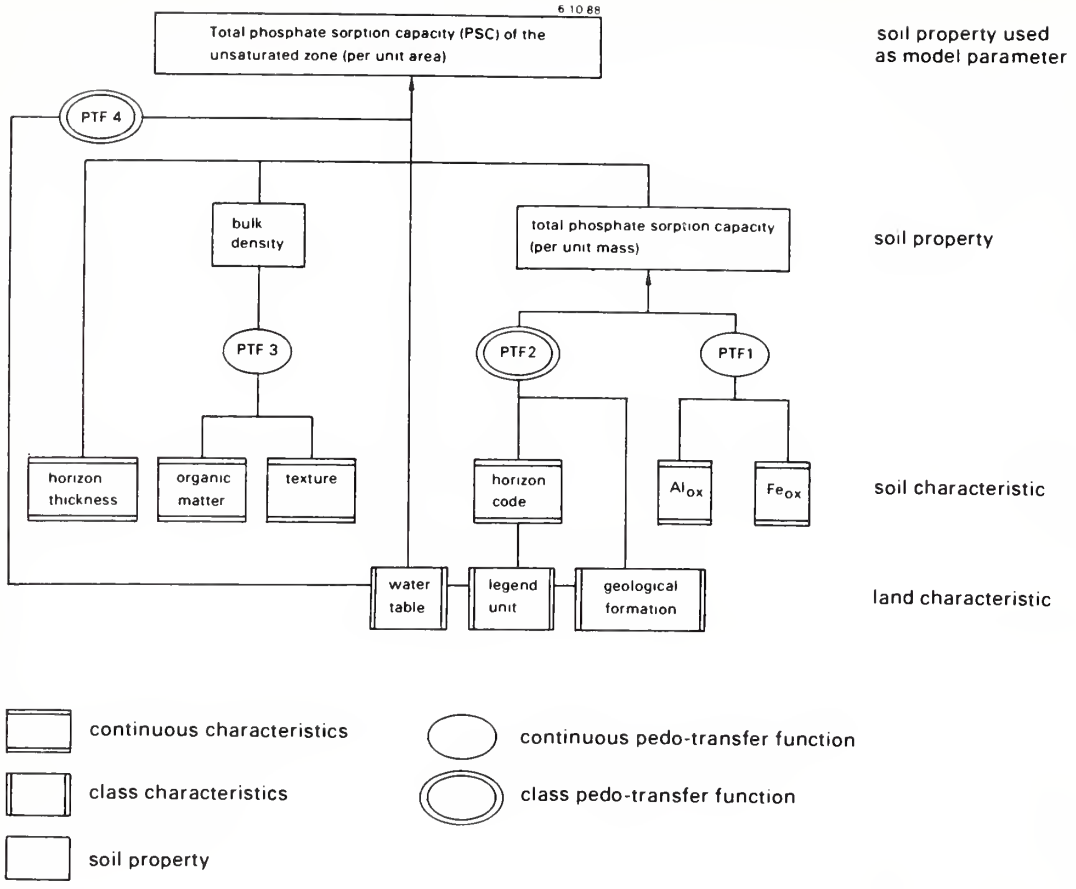


Figure 2.
Flow diagram showing the relations between model parameter and soil survey data for the total phosphate sorption capacity. (For definitions see table 2).

particular soil property, that has a variability which can primarily be attributed to the spatial properties of the feature itself and not to sampling or measurement errors. The first six aspects, involved with obtaining representative samples, will be discussed briefly.

Selection of Method

Once the need to obtain a measurement has been defined, a method needs to be selected. Soil conditions and operational considerations should play an important role in method selection. Unfortunately, this is often not the case. Substantial variability can originate from using the wrong method for the conditions or by applying, for instance, a complicated technical procedure with relatively untrained personnel. Some methods use complicated calculation procedures including substantial error even when applied professionally (see Vachaud 1982). Others yield data directly. A qualitative review of sixteen methods for the measurement of hydraulic conductivity of saturated soil (K_{sat}) and of eleven methods for the measurement of K of unsaturated soil (K_{unsat}) was presented by Bouma (1983), emphasizing aspects such as: (1) time needed for preparation, execution and calculations; (2) costs of personnel and materials; (3) complexity; and (4) accuracy.

Sample Dimensions

Many measurement procedures use standard sample sizes, because of fixed dimensions of sampling cylinders or of equipment being used. For example, sampling cylinders with a fixed volume of 100 cm^3 have been used extensively in different laboratories. Equipment, such as the double-ring

infiltrometer or the air-permeameter, comes in standard sizes. There is good justification to vary sample size as a function of soil structure, as a means to reduce variability among replicate measurements (Bouma 1983). Soil structure descriptions as made during soil survey can be used to tentatively define representative elementary volumes of samples (REV's), which are the smallest sample volumes that can represent a given soil horizon by producing an unbiased population of data. To do so, the elementary units of soil structure (ELUS) must be distinguished. These are individual sand grains in sandy soils and natural aggregates ("peds") in aggregated soils. Peds can vary in size up to several liters in very coarse, prismatic subsoil structures. Samples should contain at least 20 ELUS to be representative.

Sample Locations Inside the Soil

Sampling at regular depth intervals is often applied with good results in relatively homogeneous soils with weakly developed soil horizons. When clear soil horizons exist, however, it is preferable to sample by horizon (Peterson & Calvin 1965). A sample containing fragments of two adjacent, and as such quite different, soil horizons, will yield physical data that are difficult to interpret. However, it should be realized that pedological horizons as distinguished in soil survey are not always good "carriers" of data that are relevant to the particular measurement being made because some pedological distinctions may be irrelevant in this context. In general, it is advisable to make a soil description before making measurements. Samples should preferably be taken in soil layers with a more or less homogeneous structure, as observed in the field.

Number of Replicates

Having selected the proper method, sample dimension and sampling location in the soil profile, the investigator is faced with the question of how many replicate samples to take. The question is discussed in detail in any statistical handbook to which the reader is referred (e.g., Snedecor and Cochran 1967 Becket and Webster 1971). The number of samples is a function of the required accuracy: the latter will be higher as the number of samples increase. Graphs have been developed that allow rapid estimation of the number of samples as a function of accuracy obtained (e.g. Wilding and Drees 1983). The trade-off, of course, is cost.

Time of Sampling

Many soil structural features change during the various seasons of the year. For example, water extraction by evapotranspiration in the growing season will result in the cracking of clay soils and in semi-irreversible drying of some peaty and sandy soils. Wetting, during autumn, winter and early spring, will result in the swelling of clay soils and to some extent peaty soils. Swelling and shrinkage processes are, among other factors, a function of the rate of wetting and of the varying electrolyte content of the soil solution. Usually, these long-term processes cannot be compressed into a very short period. For example, when measuring the saturated hydraulic conductivity (K_{sat}), the soil should have been very wet or saturated for several weeks before measurement. Measuring K_{sat} in an initially dry soil is meaningless because short-term rapid swelling will result in different porosity patterns than natural, long-term swelling. In the Netherlands, K_{sat} values or moisture retention curves of clay soils, are only measured on samples that have been taken in the period February to March, in early spring when soils have been naturally wet for several months.

Irregular Flow Patterns

Flow theory assumes the presence of homogeneous, isotropic soil. Calculations of hydraulic conductivities, sorption values, infiltration rates and moisture retention characteristics are based on this assumption. Flow patterns in soil may, however, be quite irregular due to uneven infiltration at the soil surface or to bypass flow which is the movement of free water along air-filled macropores in an unsaturated soil matrix (White 1985). Bypass flow may have a major effect on water movement in many soils. Van Stiphout et al. (1987) studied water infiltration into a shallow clay soil with vertical cracks to a depth of 60 cm below the surface, covering a sandy loam deposit without macropores. Worm channels occurred to a depth of 120 cm below the surface. After two showers, wetting occurred on the soil surface and from the bottom of the cracks at 60

cm depth and from the bottom of the worm channels at 120 cm depth. A simulation model, assuming the presence of homogeneous soil predicted wetting to only 4 cm. Very prominent bypass flow was also measured in a heavy clay soil by Bouma and de Laat (1981). Soil morphological analyses can be made to define flow patterns which can then be used as physical boundary conditions for the flow system. A recent study in sandy soils indicates that preferential flow may occur in more soil types than is generally assumed (Hendrickx et al. 1988).

Use of Soil Maps and (Geo)statistics when Collecting Soil Data

Sampling Within Soil Mapping Units

So far, the required number of replicates has been discussed, but not their location in the field. Replicate samples for physical data have to be taken in soil pits. Samples can be taken in one soil profile pit by sampling the four profile walls. However, it may be preferable to take individual samples farther apart if an area of land is to be characterized. This implies, of course, that several soil pits need to be dug. Various sampling schemes for obtaining soil data in the field were recently reviewed by Wilding and Drees (1983). It is advantageous to have a good knowledge of soil and landscape conditions when choosing sampling sites in an area. Random sampling is suitable only when soil differences are not evident or not relevant to the property to be measured. In other words, when a soil map is available, it is advisable to sample at random *within soil mapping units* as defined on the soil map. Sampling points may be located in line-transects or grids to allow easier location of sampling points in the field (e.g. De Gruijter and Marsman 1985). Different soil mapping units are distinguished on the basis of pedological criteria, e.g., soil texture, organic-matter content, occurrence of specific soil horizons etc. Each mapping unit is characterized by a so-called "representative soil profile" which is defined by the soil surveyor on the basis of experience and has to be accepted as such. This presents problems as will be discussed in the following subchapter. Accepting the "representative soil profile" for the moment, it is clear that interpretations of soil maps always deal with practical applications, such as using physical and chemical properties related to agricultural or environmental problems. Pedological criteria, being used to define the mapping units of the soil map, are therefore not necessarily relevant for these applications. This aspect was analyzed by Wösten et al. (1985) and Breeuwsma et al. (1986). They tested the relevance of differences among all distinguished soil horizons in "representative profiles" for particular applications. Wösten et al. (1985) showed that of nine distinguished major soil horizons in an area of 650 ha with sandy soils in The Netherlands, only five horizons had significantly different hydraulic properties in terms of hydraulic conductivity and moisture retention. For a sample test area of 125 ha containing 110 delineated areas on the soil map, only 41 areas could be distinguished which had significantly different hydraulic properties. The map showing these 41 areas was called a "simulation map" as it contained for each unit basic soil physical information for simulations of the soil water regime. Breeuwsma et al. (1986) demonstrated an identical reduction of mapping units when considering the cation exchange capacity and the phosphate sorption capacity. The question may be raised as to the reliability of such derived maps. The reliability of their "simulation maps" was tested by Wösten et al. (1985) by making 60 random test borings which indicated an accuracy of about 80%. This is considered to be adequate for their particular study. However, because of the subjective character of the "representative profiles" this result does not necessarily relate to other studies.

Using Point Data and Geostatistical Interpolation Techniques

As discussed above, the definition of "representative profiles" has a rather subjective character. Even though results obtained may sometimes allow reliable predictions, there still is a need for procedures that are completely quantitative and reproducible. One procedure is to use calculations for exact point data only, to be followed by interpolation using geostatistical techniques. In recent years, there have been strong advances in geostatistics using the theory of regionalized variables (Journel and Hybregts 1978). These regionalized variables have values that are related in some way to their position. Basically, geostatistical theory states that observations which are located closely

together are likely to have a higher probability of resembling one another than observations that are farther apart. This phenomenon can be mathematically expressed by autocorrelation, by a semi-variogram or by intrinsic random functions (e.g., Nielsen and Bouma, 1985). A known semi-variance can be used to predict values at unmeasured locations by interpolation using the kriging technique, which is being discussed by many others during this conference. Bregt et al. (1987) used kriging to predict simulation maps for the study area described by Wösten et al. (1985) and obtained maps with the same accuracy as the simulation maps from the soil map. However, with kriging, a specific measure for accuracy can be obtained for any point estimate and this is not possible when using "representative profiles"; the use of geostatistics is therefore very attractive. Geostatistics can also be used effectively for optimizing sampling distances. For example, Vieira et al. (1981) and McBratney and Webster (1983) have shown that the use of geostatistics can strongly reduce the number of required replicate samples while still attaining an acceptable degree of accuracy.

RECENT APPLICATIONS OF DATA BASES IN LEACHING MODELS

The possible role of data bases in the application of leaching models may be illustrated with a few recent studies on the leaching of nitrate and phosphate. The leaching of nitrate is increasingly studied as a result of high nitrate levels in groundwater caused by the intensive use of fertilizers and manure. The leaching of phosphate is entirely caused by the over use of manure and therefore restricted to areas with intensive livestock farming as present in, for example, the Netherlands. The studies considered involve applications outside the US and usage on a regional scale. However, the same approach may be used on a field scale as far as soil data are concerned.

Nitrate

The first example deals with the use of computerized soil and climate data in the UK (Jones et al. 1987). The soil map scale 1:250,000 was used to derive the dominant soil association in each 5 km^2 area and each association assigned to a leaching risk class on the basis of permeability and parent material. Excess winter rainfall was calculated for each 5-km grid using average monthly rainfall for 970 stations in England and Wales. Potential nitrate losses from land under five crops and typical applications of N fertilizer were estimated using data from monitoring studies (Robson et al. 1987). Thematic maps showing the nitrate losses (by leaching) for a particular monoculture in each grid were produced by a data base management system. This example illustrates how data bases may be applied in combination with field measurements using an empirical model. Empirical models may be attractive in describing present situations, but results are difficult to extrapolate to other conditions with regard to soil types, water regimes, cropping systems and fertilizing practices. This extrapolation requires a process-based model, for example the Addiscott model used to test the leaching risk classification applied in the above study (Carter et al. 1987).

The second example differs from the first in using a process-oriented model, actual data on land use and more detailed information on soil types (not only the dominant type). A regional nitrate leaching model (RENLEM) is being developed to forecast the long-term effects of fertilization on the leaching of nitrate from soils on a regional scale (De Vries et al. 1987). The N transformation processes are modeled in a simple way. The N status of the soil is assumed to be at equilibrium and nitrification is considered to be completed within the summer period. Standardized rates are used for volatilization and uptake by the crop. Water flux and moisture content are calculated by a separate hydrologic model. Pedo-transfer functions are used to calculate moisture retention and hydraulic conductivity curves as described by Wösten et al. (1985). Denitrification is the major soil process in controlling the leaching of nitrate in soils with shallow water tables which frequently occur in the Netherlands. The modeling of the N processes, therefore, was focussed on the modeling of the denitrification process. Three ways of modeling the denitrification rate have been used:

(1) A first-order reaction and a rate constant controlled by pH, temperature and moisture content.
(2) A first order reaction and a rate constant related to soil type and water-table level by a pedo-transfer function.
(3) A constant fraction of the net input related to soil type and water-table level by another pedo-transfer function.

The denitrification rate constant was calibrated using literature data, field data and lysimeter data, respectively. The pedo-transfer function derived by the third method is shown in table 3 to illustrate the significant effect of the water regime on the denitrification rate in sandy soils with shallow water tables.

The model is now applied to water supply areas to calculate the nitrate load to the phreatic surface for each combination of soil type (including the water-table class) and land use for various N input scenarios. The areal distribution of soil types over land use is derived from an overlay of the soil map and the topographic map scale 1:25,000 by determining the combinations of soil types and land use at the intersection points of a grid network. The total N load to the phreatic surface in a particular area is calculated by taking the area-weighted average of the N loads of all combinations of soil type and cropping systems. A digitalized soil data base allows easy graphical presentation of the results if a monoculture of typical crops and standardized rates of application are assumed for particular soils. Figure 3 presents an interpretive map derived from the soil map in this way to show the areas where the soil is vulnerable to leaching of nitrate (De Vries et al. 1987). The areas with high nitrate concentrations (above 100 mg/l) are primarily forested sandy soils on *ice-pushed ridges*. The relatively high concentrations are caused by high deposition rates (50 kg N/ha) and a low precipitation surplus (160 mm/yr. for coniferous forests). Moreover, the denitrification rate is low because of the deep water-table level (class VII*). The areas with low concentrations of nitrate to the phreatic surface are mainly clay soils with shallow water-tables (class III and V). They were assumed to be used as grassland receiving 490 kg N/ha.

Phosphate

To our knowledge, the third case study presented here is one of the most advanced applications of data bases in leaching models. The first part of this study involves the application of a regional phosphate transport model (REPTRAM) in forecasting the breakthrough of phosphate in intensive livestock areas (Breeuwsma and Schoumans 1987). Soil boundaries were derived from the soil map scale 1:50,000, soil characteristics from the Soil Information System (BIS) and pedo-transfer functions from additional research (Schoumans et al. 1987). Data on land use were derived from the topographic map and the distribution of soil types (and water-table classes) over grassland and arable land obtained as described for nitrate. The present amounts of phosphate in the soil were estimated from measurements of the natural quantities and from statistical data on the production of manures in the past using the available data for agricultural areas of about 10,000 to 100,000 ha of cropland. A simple model assuming a step-front movement was used to calculate the area of soils saturated with phosphate at a given depth as a function of various input scenarios (figure 4).

The calculations were made for maize because application rates for this crop are often very high. The model calculations indicate a significant part of the maize area studied (about 100,000 ha) to be already saturated to the mean highest water level. This situation could result in leaching to surface waters by shallow groundwater flow of amounts of phosphate that may well exceed the phosphate load by surface runoff in the flat areas under investigation. Studies such as these, together with field measurements, have recently led to regulations that require the mapping of risk areas and identification of phosphate-saturated fields.

Table 3.
Pedo-transfer function relating the denitrification fraction to the water-table class in sandy soils (after Steenvoorden 1984).

Water-table class[1)]	Mean highest water level	Mean lowest water level	denitrification fraction[2)]
	-cm below surface-		
II	<40	50-80	0.96
III	<40	80-120	0.90
IV	>40	80-120	0.78
V	<40	>120	0.85
VI	40-80	>120	0.59
VII	>80	>160	0.17
VII*	>140	>200	0.00

1) as defined by the Netherlands Soil Survey Institute
2) Fraction of the net N input to the phreatic surface at water-table class VII* which is denitrified

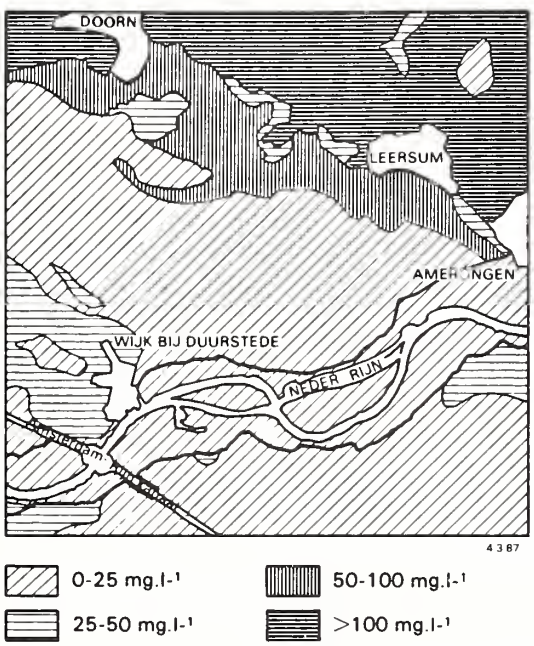


Figure 3.
Map of nitrate concentration classes of phreatic water derived from the soil map by using a regional nitrate leaching model (De Vries et al 1987).

The second part of this study deals with the application of the data bases and models in the mapping of risk areas. The major differences with the first part are the use of digitized data bases, remote sensing information and more detailed manure data. Data on land use (mainly grassland and maize) are derived from remote sensing images. This is because of the actual information on land use and crops that may be obtained and because of the digital information storage. Information from the Landsat Thematic Mapper with pixels of 25 m^2 yielded excellent identification of grassland and maize fields in the relevant agricultural areas (Van der Laan et al. 1987). The basic unit of storage of the digitized soil map 1:50,000 is adjusted to that of the land use data base. Application rates of manure are based on production data calculated from annual records of livestock which are made by the Ministry of Agriculture. For reasons of privacy,

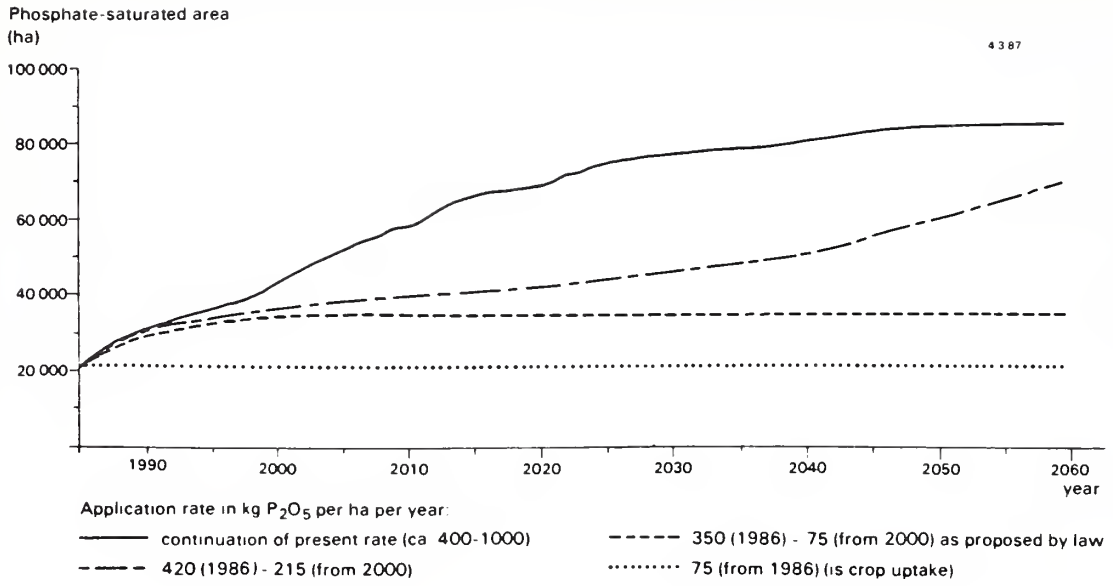


Figure 4.
Long-term development of the area of phosphate-saturated maize land in The Netherlands for four scenarios of phosphate application (saturation depth: mean highest water level) (Breeuwsma and Schoumans 1987).

production figures must be lumped on a grid level of, probably, 2 km^2. Within each grid, standardized rates of application are used for grassland and maize. The data on soils, land use and application rates are then combined to assess the area saturated with phosphate at a given depth for 2-km grids. The saturation is expressed in terms of the percentage of phosphate-saturated soils and each grid allocated to a risk class. Figure 5 illustrates the model output for an area of 50,000 ha assuming annual application rates of 800 and 300 kg of P_2O_5 per ha for maize and grassland, respectively, for a period of 16 years.

The use of average data on inputs per grid and model parameters per soil type, allows a qualitative risk assessment. However, the same approach can be used to obtain a quantitative assessment by using a stochastic model that accounts for the distribution of data within each grid and soil type. Further research will address this topic as well as the validation of the model on a regional scale and the leaching of phosphate to surface waters.

CONCLUSIONS

(1) Applications of models which predict leaching of pollutants from the root zone on a regional scale need data bases on weather, soils, land use and application of chemicals.
(2) Soil survey data may provide essential and readily-available information on soil properties and land qualities used as model parameters when

    (a) model parameters are related to soil and land characteristics (by so-called pedo-transfer functions);

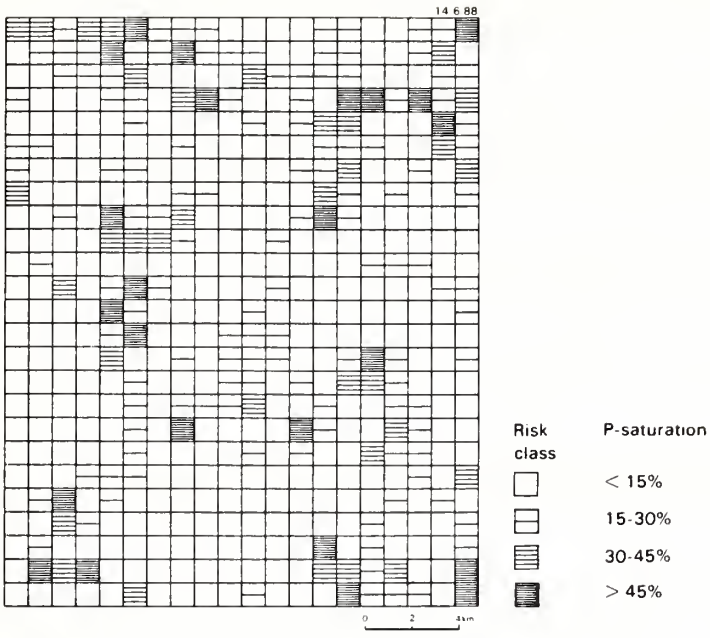


Figure 5.
Grid map of risk classes for phosphate saturation derived from digital data bases by using a regional phosphate transport model (Thunnissen and Schoumans, in press).

(b) parameter values (including their distribution) are representative of a particular area as shown on the soil map made by the soil surveyor or by geostatistical methods;
(c) the information on soil boundaries and characteristics is stored in digitalized data bases.

To benefit from soil survey data modelists need to pay more attention to pedo-transfer functions and soil surveyors to the use of soil maps and/or point data in deriving representative spatial data on model parameters.

(3) Remote sensing provides a powerful tool in collecting actual data bases on land use in a digitalized form.

LITERATURE CITED

Abbott, M.B., J.C. Bathurst, J.A. Cunge, and others. 1986. An introduction to the European hydrological system - Systeme Hydrologique Europeen "SHE", 1: History and philosophy of a physically-based, distributed modelling system. J. Hydrol. 87:45-59.

Beckett, P.H.T., and R. Webster. 1971. Soil variability; A review. Soils Fert. 34:1-15.

Bouma, J., 1983. Use of soil survey data to select measurement techniques for hydraulic conductivity. Agric. Water Manage. 6:177-190.

Bouma, J., and P.J.M. de Laat. 1981. Estimation of the moisture supply capacity of some swelling clay soils in the Netherlands. J. Hydrol. 49:247-259.

Bouma, J., and H.A.J. van Lanen. 1987. Transfer functions and threshold values: From soil characteristics to land qualities. In Quantified Land Evaluation Procedures K.J. Beek, P.

Burrough and D.E. McCormack (Eds.), ITC Publication no. 6, Enschede, Netherlands. 106-110.

Bouma, J., H.A.J. van Lanen, A. Breeuwsma, and others. 1986. Soil survey data needs when studying modern land use problems. Soil Use and Management 2:125-130.

Breeuwsma, A., J.H.M. Wösten, J.J. Vleeshouwer, and others. 1986. Derivation of land qualities to assess environmental problems from soil surveys. Soil Sci. Soc. Am. J. 50. 1:186-190.

Breeuwsma, A., and O.F. Schoumans. 1987. Forecasting phosphate leaching from soils on a regional scale. Proc. Int. Conf. on the Vulnerability of Soil and Groundwater to Pollutants. Noordwijk and Zee, 1987, The Netherlands 973-981.

Bregt, A.K., J. Bouma, and M. Jellinek. 1987. Comparison of thematic maps derived from a soil map and from kriging of point data. Geoderma 39:281-291.

Bregt, A.K., J. Denneboom, J. Stolp, and J.J. Vleeshouwer. 1986. BIS, an information system for soil survey and research. Trans. XII Congress of the International Society of Soil Science, Hamburg, Fed. Rep. Germany, 3:1064-1065.

Carter, A.D., R.C. Palmer, and R.A. Monkhouse. 1987. Mapping the vulnerability of groundwater to pollution from agricultural practice particularly with respect to nitrate. In: Proc. Int. Conf. on Vulnerability of Soil and Groundwater to Pollutants, Noordwijk aan Zee, TNO Committee on Hydrological Research no. 38, The Hague, The Netherlands 333-342.

De Gruijter, J.J., and B.A. Marsman. 1985. Transect sampling for reliable information on mapping units. In Nielsen, D.R. and Bouma, J. (Eds.), Soil spatial variability. Proceedings of a Workshop of the ISSS and the SSSA, Las Vegas, USA, 30 Nov.-1 Dec. 1984. pp. 150-166, Pudoc, Wageningen.

Hendrickx, J.M.H., L.W. Dekker, E.J. van Zuilen and O.H. Boersma. 1988. Water and solute movement through a water repellent sand soil with grasscover. Proc. Int. Conf. on Validation of flow and transport models for the unsaturated zone, New Mexico, USA.

Jones, R.J.A., A.J. Thomasson, J.D. Robson, and A.D. Carter. 1987. The distribution of potential nitrate leaching losses from agricultural land in the UK based on computerized soil and climate data. Report of the Soil Survey and Land Research Centre, Rothamsted Experimental Station, Harpenden, UK, 16 pp.

Journel, A.G., and Ch.J. Huybregts. 1978. Mining geostatistics. Academic Press. London.

Jury, W.A. 1982. Simulation of solute transport using a transfer function. Water Resour. Res. 18:363-368.

Laan F.B. van der, O.F. Schoumans, and H.A.M. Thunnissen. 1987. Land use inventory using Landsat thematic mapper imagery to study environmental pollution. Willi Nordberg Symposium, Austria. National Aerospace Laboratory (NLR), the Netherlands, report nr. NLR MP 87063 U, 11 pp.

Lamp, J., and W. Kneib. 1981. Zur quantitativen Erfassung und Bewertung von Pedofunktionen. Mitt. Deutsch. Bodenkundl. Gesellschaft 32:695-711.

McBratney, A.B., and R. Webster. 1983. How many observations are needed for regional estimation of soil properties? Soil Science 135:177-183. Nielsen, D.R., and J. Bouma, (Eds.), 1985. Soil spatial Variability. Proc. of an ISSS-SSSA workshop, Las Vegas, USA. Pudoc, Wageningen, the Netherlands.

Nielsen, D.R., and J. Bouma. 1985. Soil spatial variability. Proceedings of an ISSS-SSSA workshop. Pudoc, Wageningen, The Netherlands.

Petersen, R.G., and L.D. Calvin. 1965. Sampling. Chapter 5 In Methods of Soil Analysis. Part 1. C.A. Black (ed.), Am. Soc. Agron. 9:54-73.

Robson, J.D., P.J. Loveland, and A.D. Carter. 1987. Field measurement of soil nitrate status and leaching losses from farmland. Report of the Soil Survey and Land Research Centre, Rothamsted Experimental Station, Harpenden, UK, 56 pp.

Schoumans, O.F., A. Breeuwsma, and W. de Vries. 1987. Use of soil survey information for assessing the phosphate sorption capacity of heavily manured soils. Proc. Int. Conf. on the Vulnerability of Soil and Groundwater to Pollutants. Noordwijk aan Zee, 1987, The Netherlands 1079-1088.

Snedecor, G.W., and W.G. Cochran. 1967. Statistical Methods (6th Ed.), Iowa State Univ. Press. Ames. Iowa. 593 pp.

Steenvoorden, J.H.A.M. 1984. Effects of changes in water management on water quality (in Dutch). Report no. 1554. Institute for Land and Water Management (ICW), Wageningen, The Netherlands, 18 pp.

Stiphout, T.P.J. van, H.A.J. van Lanen, O.H. Boersma, and J. Bouma. 1987. The effect of bypass flow and internal catchment of rain on the water regime in a clay loam grassland soil. J. Hydrol. 95:1-11.

Vachaud, G. 1982. Soil physics research and water management. Whither Soil Research. Proc. ISSS Congress, India 32-59.

Vieira, S.R., D.R. Nielsen, and J.W. Bigger. 1981. Spatial variability of field measured infiltration rate. Soil Sci. Soc. Am. J. 45:1040-1048.

De Vries, W., O.F. Schoumans, J.F. Kragt, and A. Breeuwsma. 1989. Use of models and soil survey information in assessing regional water quality. G. Jousma et al. (eds.) Groundwater contamination: Use of models in decision-making, Kluwer Academie Publ.: 419-432.

De Vries, W., J.F. Kragt, and A. Breeuwsma. 1987. The use of soil maps in predicting nitrate leaching with a regional transport model. Proc. Int. Conf. on the Vulnerability of Soil and Groundwater to pollutants. Noordwijk aan Zee, 1987, The Netherlands 491-498.

White, R.E. 1985. The influence of macropores on the transport of dissolved and suspended matter through soil. In B.A. Stewart (Ed.) Advances in Soil Science. 3:95-121.

Wilding, L.P., and L.R. Drees. 1983. Spatial variability and pedology. In Pedogenesis and Soil Taxonomy. L.P. Wilding, N.E. Schmeck, and G.F. Hall (eds.), Developments in Soil Science IIA-Elsevier, Amsterdam 83-113.

Wösten, J.H.M., J. Bouma, and G.H. Stoffelsen. 1985. Use of soil survey data for regional soil water simulation models. Soil Sci. Soc. Am. J. 49:1238-1244.

DATA MANAGEMENT FOR WATER-QUALITY MODELING DEVELOPMENT AND USE

Alan M. Lumb[1], Robert F. Carsel[2], John L. Kittle, Jr.[3]

ABSTRACT

Improved usage of greater volumes of data is a key to improved water-quality analysis. Types of data used for water-quality modeling include meteorologic, hydrologic, constituent loads and concentrations, vegetal cover, land use, soils aquifer properties, topography, chemical properties, biologic, and toxicity. Several data sources and data-management tools are discussed in the context of water-quality modeling.

INTRODUCTION

A key to better protection of our water resources and effective use of water-quality models is the use of greater volumes of more accurate data. Generally, the understanding of hydraulic phenomena and the physical, chemical, and biological processes derived from laboratory and field experiments exceeds our ability to collect and manage sufficient data to define boundary conditions and to fully test algorithms of the processes on large drainage basins, rivers, and lakes. Data on the spatial distribution of soil characteristics in a watershed are inadequate relative to our understanding of how those characteristics affect infiltration. The understanding of transpiration processes exceeds our ability to manage spatially distributed data on vegetation type, slope, aspect, and soils. Our understanding of the movement of water and solutes in the unsaturated and saturated zones far exceeds our ability to define the spatial distribution of the porosity, hydraulic conductivity, and organic content of soils. Precipitation, one of the main driving forces in water-quality models, is only measured at a few locations in or near a watershed. Yet, spatial and temporal variation of storm precipitation is substantial. One of the largest components of model error is precipitation input.

Several technologies have emerged and are being explored to provide data for water-quality modeling with a much higher spatial and temporal resolution than previously available: programmable data loggers are becoming inexpensive and can store large volumes of data from many sensors; data can be relayed from satellites as it is being observed; resolution of satellite data from multiple wave lengths is being reduced; and cartographic data on land features and topography are rapidly being digitized. The next generation of radar will provide hourly estimates of rainfall on a 2-mile grid.

The more efficiently we are able to manage data from these new sources, the greater the potential for accurate water-quality modeling. Researchers will be able to test new algorithms on a variety of watersheds. Water managers will be able to screen a wide range of management scenarios. Effective data-base management integrated with flexible, modular software can be a major tool for technology transfer from researchers to water managers.

---

[1]Alan M. Lumb, Hydrologist, U.S. Geological Survey, Reston, VA.

[2]Robert F. Carsel, Environmental Scientist, U.S. Environmental Protection Agency, Athens, GA.

[3]John L. Kittle, Jr., Senior Systems Analyst, Aqua-Terra Consultants, Atlanta, GA.

Some of the problems encountered with the new types of data as well as some of the existing data are the lack of standard formats and definitions of parameters. Each time a new format is used, software has to be modified. Each data logger has its own format as does each Government agency. Although much effort is being made to develop the standards, the problem persists. When working with large volumes of data, quality control of that data can become a major task. The researcher or water manager no longer is able to view all data. Much of the data are transferred from the data sensor to the data base without being examined in a listing or a graph. For time-series data, periods of missing record frequently occur. In many cases estimates for those periods are needed before a water-quality model can be used. Climate data commonly have periods of missing record, and, in general, resources are not sufficient for researchers or managers most qualified to review and adjust the record.

The major problem with using large volumes of data for water-quality models is the lack of tools for efficient use of the data. Software to efficiently manage the data can be expensive to develop or purchase for each site using the model. Most available data-management systems were not designed for large volumes of time-series data. Many data-management systems do not function effectively with water-quality models.

This paper discusses data sources and data management. A brief summary of data sources will be presented to illustrate the breadth of the problem and to identify the range in types of data that can be useful to water-quality modeling. The section on data management will identify some of the issues and describe one approach to managing data for water-quality modeling.

DATA TYPES AND SOURCES

Many types of data are used directly or indirectly by water-quality models. Rainfall and air temperature data are usually direct input to a model. Data on channel cross sections commonly are used to determine flow properties that then are used for direct input to a model. To begin to address the data-management needs, these types of data need to be identified and categorized. This section identifies the types of data and the most common sources for that data.

Meteorologic

The atmospheric data frequently needed for water-quality modeling include precipitation, temperature, solar radiation, humidity, wind velocity, wind direction, snowfall, and pan evaporation. The major source of data is the National Climatic Center in Asheville, North Carolina (Hatch 1983), although additional data are available from other Federal agencies, water-management districts, research stations, public utilities, and universities.

Water-Quality Loads and Concentrations

The largest sources of water-quality data are the Environmental Protection Agency (EPA) and the U.S. Geological Survey (USGS) (Edwards 1980). State and local governments also collect a large amount of water-quality data, much of which is provided to the EPA. Quality-control procedures vary widely with the collected data. The EPA maintains the STORET data-base system that serves as a national repository for these data. On that system each measured constituent is identified by a code and the location, date, and time of collection.

Vegetal Cover and Land Use

There are many sources of data including maps, tables, and digital representations. Digital land-use data are available from the USGS National Cartographic Information Center (NCIC) (1985). Large volumes of remotely sensed data are available. The most common are the

LANDSAT thematic mapper imagery (TM) and the multispectral scanner imagery (MSS). LANDSAT data require much processing to produce vegetal cover for a selected area.

Soils

Soils data are available from surveys conducted by the Soil Conservation Service (SCS) for the majority of counties within the United States. These surveys contain a description of the soils present within the county and information on the use, texture, engineering, and areal distribution of each soil. Other sources include Soil Survey Information Reports. Contained in these reports are selected morphological, textural, water retention, engineering, and use information. These data bases are not automated and usually require considerable effort to obtain desired soils data.

There are three soils data bases that are automated: (1) the Soils Information Retrieval Information System (SIRS) (U.S. Department of Agriculture [USDA] 1985), (2) the National Resources Inventory (NRI) (USDA 1982), and (3) the Data Base Analysis for Pesticide Evaluations (DBAPE) (1988). The first two are SCS data bases/products, whereas the last is an agricultural subset of SIRS developed by EPA.

The information provided in these data bases consists of land and water use, soil characteristics, land-management practices, cropping information, and geographical distribution. SIRS and NRI data bases are quite large and are implemented on mainframe computers. The DBAPE data base is designed to run on microcomputers.

Aquifer Boundaries and Properties

The major source of information for aquifer information is the USGS WATSTORE system (Baker and Foulk 1984). The information contained in WATSTORE can provide much of the information needed for modeling solute transport including the depth of the aquifer, aquifer material, location, and hydraulic properties. WATSTORE is large and has several distinct data bases and retrieval procedures that can be difficult to learn. Selected USGS water-resources publications also contain information on aquifer boundaries and properties.

Topography

Topographic data includes digital data that describes the land surface, location of drainage divides and channels, and channel cross sections. NCIC (1985) has digital data on hydrologic features and grids of land-surface elevations at several scales and resolutions. The EPA maintains a river-reach file of all the major rivers and tributaries in the United States. Channel cross-section data exist in many formats in many locations; the major source is from studies for flood-insurance reports for the Federal Emergency Management Agency.

Chemical

Chemical properties for organic compounds have not been organized into any central data base (there are efforts to establish a centralized system). Characteristics for selected organic and inorganic constituents are obtained from a variety of sources including chemical handbooks (e.g., Merck Index [Windholz 1983], and agricultural farm chemical handbooks [Hartley and Kidd 1983]), data bases contained in models (e.g., MINTEQ and PRZM), and agency (e.g., Federal and state) reports. The data contained in these sources are highly variable. Some chemicals contain most all common physico-chemical characteristics (e.g., solubility, octanol water partitioning, vapor pressure, and Henry's law constant), whereas other characteristics (e.g., adsorption kinetics and degradation rates) are not readily available.

Biological

Biological data are similar to the chemical data in that they are not organized into a centralized data base. The U.S. Department of Agriculture maintains an extensive data base on pesticide transformation for several hundred chemicals (Nash 1988). Most data for biological characteristics are contained in selected literature citations.

Toxicity

The toxicological information for chemicals is maintained by the National Institutes of Health (NIH) (Gleason 1969). They retain information on acute, chronic, and mutagenic characteristics. Other sources are chemical handbooks such as Merck's and literature citations. Most of the NIH data are in computer format.

DATA MANAGEMENT

Needs and Requirements

Data management needs for water-quality modeling can be classified by the user--research, development, or applications; and by task--preprocessing, modeling, or postprocessing. Each of the user groups performs each task, but the emphasis will be different. The applications group is most interested in postprocessing, whereas the development group is most interested in model calibration and validation.

Preprocessing

Preprocessing needs include obtaining a data base from various sources and integrating it in a data management system that can be read by the water-quality model. Although the task is usually straightforward, it is complicated by the range of formats that are used and the schemes used to flag or skip missing data. Preprocessing also must identify and update bad data through the use of graphics and statistical analysis tools.

Preprocessing also includes the need to change some of the data as part of water-quality model calibration or the development of different simulation scenarios. The user needs to be able to quickly and logically locate and change part of the data.

Modeling

Water-quality modeling commonly involves the processing of large volumes of data as input and output. The data must be located on the disk quickly by using a few names or numbers that are input to the model. The transfer of data from the disk to the buffers in the model needs to be fast with as little overhead processing as feasible. Transfers of computed data from the buffers in the model to the disk also needs to be fast with little overhead. Again the names and numbers needed to locate the output data on the disk must be few and specified as input to the model.

Postprocessing

Postprocessing has many of the same needs as preprocessing, including statistical analyses and tabular and graphic presentations of the water-quality modeling results. The data-management system needs to support a flexible and logical means to select and present the data. Various types of summary tables are also needed.

Calibration aids, model parameter optimization techniques, and sensitivity analysis are needs that could be listed under the modeling or postprocessing task. Data management systems must meet these needs.

Special Requirements

User groups in water-quality modeling are found in a wide variety of government agencies, universities, and private firms. The need to share models and data is great. Water-quality models are moved from research groups to groups developing the model for a particular drainage basin. Once the model has been developed (calibrated, verified, and validated), it may be moved to another users group for application. All three groups may be from different agencies, and all three groups may use different computer systems. Thus, the software and data base need to be portable from one system to another. Operating systems, compilers, and graphics libraries are always required, but they can be selected to minimize the portability problem. Use of proprietary software for data management and interactive display needs to be minimal because they can present major problems because of costs, obsolescence, limited availability on different types of computer systems, or purchase restrictions.

Data Categories

Data for water-quality models can be placed in one of three categories:

(1) basin schematics,
(2) tables of properties and options, or
(3) time series.

Most all water-quality models simulate conditions in a lake, a section of river, an aquifer, a land surface, or an entire river basin. Simulations may occur in space, in time, or in both. Basin schematics are needed for modeling spatially distributed variables to define the association of the spatial elements. The schematic may be simply a string of reaches along a river or may be complex, such as land segments discharging to river reaches, with reservoirs, water withdrawals, and wastewater discharges. In some water-quality modeling, the schematic is a grid of elements or nodes frequently used in groundwater or estuarine modeling. Schematics define direction of flow, feed-back loops, and order of computations.

Tables of properties include static data and state variables for points, lengths, areas, or volumes. For an agricultural field, it may include size, slope, crop type, application rates, and coefficients for infiltration equations. For a channel cross section it may be a table of flow properties by elevation. For a lake it may be a table of volume, surface area, and discharge by elevation.

Time-series data are input to or output from a water-quality model and may represent fluxes, depths, or volumes over points, lines, or areas. Precipitation on a cornfield, wastewater discharge to a stream, or water-surface elevation of a lake are examples. Each time series must be associated with a point, line, or area for proper handling by the model.

Thus, the basin schematic is used to associate the tables and time series with processes in the water-quality model. The basin schematic is the key to logically managing the data.

Tools

There are many concepts and tools for managing data (Date 1985). Basically, the tool is a computerized record-keeping system to add, insert, retrieve, update, delete, and remove data from the system. Systems are traditionally placed in one of four categories: relational, inverted list, hierarchic, and network. Relational data-base systems have become the most popular in recent years and many proprietary systems are available for any hardware system. They cost from a few hundred dollars to over fifty thousand dollars. Many different relational data-base systems are used in government agencies. In relational systems all the data are seen as tables that can be related in any number of ways. In hierarchic systems the data is represented to the user as a tree structure, much like a drainage basin with many levels of tributaries. Network systems represent

the data in a variety of ways other than a tree structure, which may include loops. Inverted lists use secondary keys as well as primary keys in structuring the data.

Data-base management systems require some effort to define all the data elements and how they relate; but once done for a system, the retrievals and updates require little effort. The drawbacks are the cost, portability, and interfaces to water-quality models. Although many systems allow access by water-quality models written in FORTRAN, BASIC, C, or Pascal, the effort is not always easy or efficient.

The option to selecting and using an existing data-base system is to write a system using a standard compiler, such as FORTRAN. The advantages of portability and retrieval speed are offset by the development cost. Tradeoffs are difficult to assess for water-quality modeling. Some simple models may most appropriately use an inexpensive data-base management system or spread-sheet software for a personal computer. More comprehensive water-quality models with high volumes of data input may most appropriately use a specially designed data base.

Two basic approaches have been used in specially designed software. One approach uses the operating system and the hierarchy of directories and files on the disk. In these cases file names are used to locate the data and the files have specific formats for the data. Examples include the geographic information system ARC-INFO; a package developed at Stanford and reported in EOS (Smith and Clauer 1986); a system developed at the Goddard Space Flight Center (Treinish 1984); and the DDS system developed at the Hydrologic Engineering Center (1983). The other approach is to store all the data in one file using a system of pointers to locate data. The Watershed Data Management system is an example of that approach (Lumb and Kittle 1985). The first approach may be easier to initially implement but can have some problems with portability since different operating systems use different conventions for directories and files. Retrievals from the first approach are usually not as fast because files have to be continually opened and closed, a process that carries notice-able overhead. Also, operating systems have an upper limit to the number of files that can be open at any one time.

The Watershed Data Management system will be described in the next section as an example of the approach chosen by the authors for use with a variety of hydrologic and water-quality models.

WATERSHED DATA MANAGEMENT SYSTEM

Purpose

The Watershed Data Management (WDM) file and the associated data-management program ANNIE/WDM are designed to be a standard system to store, update, and retrieve time-series, drainage network, and basin characteristics data for hydrologic, hydraulic, and water-quality models. The system has been developed to overcome the following problems encountered in research, development, and application of water-quality models.

The hydrologic modeler usually has to learn a different data storage and retrieval system for each model used. When the output from one model is used as input to another model, the user has to write a computer program to convert the data. Much time and many resources are wasted on learning new systems and managing data.

Model developers and users have both identified inadequacies with the current storage and retrieval systems used with models. Some systems cannot store much data. Other systems store data inefficiently or the retrievals are too slow. Many systems are difficult to use with other programs and are difficult to update for data corrections.

Government agencies, universities, and other groups are looking to the concepts of expert systems to provide better user applications of water-quality models. To provide the information for the user/computer interactions, a comprehensive, well-organized data storage and retrieval system is needed.

Most water-quality models require spatial data on sizes, characteristics, and locations of watersheds, streams, reservoirs, and flow diversions. Geographic Information Systems (GIS) can be used to determine such data as drainage boundaries, drainage areas, land slopes, channel lengths, lake area,and flow paths. The GIS software writes computed data to a standard file for use by water-quality models.

Many of the current data storage and retrieval systems for water-quality models are difficult or impossible to transfer from one type of computer system to another. Occasionally the most appropriate model is not used because its data storage and retrieval system will not work on a given system or the proprietary software is not currently available.

An ad hoc group of modelers from Federal agencies met in Denver in the spring of 1984 to discuss coordination of modeling activities for water-supply forecasts. One of the recommendations of the group was the development of a common data storage and retrieval system for hydrologic and hydraulic models. Four Federal agencies have expressed strong interest, and three agencies (USGS, EPA, and SCS) have contributed resources to design and implement the WDM file structure and software utilities.

A major premise for data management for water-quality models is that data comes in groups such as 10 years of daily streamflow, a grid of elevation for a watershed, a set of coordinates for a channel cross section, or a table of hydraulic properties for a channel. All or parts of one or more group may be needed by a model. And the groups must be identified for easy and logical retrieval by the user. The groups are called data sets, and the data set identifiers are called attributes.

Both the WDM file and associated software are portable and useful on most microcomputers with a minimum 10 megabytes hard disk, as well as on super microcomputers, minicomputers, and mainframe computers. The file system is space efficient with less than 50 percent overhead for most data. A WDM file has the potential for up to 30,000 data sets per file, yet it is also efficient and useful for only a few data sets.

For identification and retrieval, a WDM file stores from one to several hundred attributes for each data set. The user may easily expand, modify, or delete all or portions of each data set.

The user of time-series data has the option to place flags on the data to indicate quality of the data, whether the data is measured or estimated, periods of missing record, and so forth. Formats for constant or variable time steps and random observations are available. Time steps from 1 second to 1 year are supported. Any string of time-series data with identical or near-identical values can be compressed on the file depending on a tolerance specified as an attribute for the data set. Pointers within a time-series data set can be set for century, decade, year, month, day, or hour as appropriate.

Storing data from or for GIS systems and for plotting maps and cross sections will be supported. Both data in the raster and vector format are needed.

Each data set has a unique data set number for pointing to the location of the data set without sequential searching. Data sets can be linked in one or more networks using a schematic data set.

Although the file structure is moderately complex, the software can make the file easy to use. The interactive software package, ANNIE, contains the WDM utilities to manage the WDM file. The

utilities store, retrieve, and update data for use in application programs. The software uses a buffer to keep the WDM records most recently used to minimize the number of disk read and write operations. Several types of record and location pointers are included and used by the software to associate data, to find data faster, to improve storage efficiency, and to dynamically expand or reduce the size of data sets.

File Structure

The WDM file is a set of unformatted, direct access records that are 512 words or 2,048 bytes in length. Formatted, sequential files are used to archive the data and transfer the data between computer systems. ANSI standard FORTRAN 77 subroutines are used to build the contents of the buffer before writing to the file and to decipher a buffer after a record is read. A set of FORTRAN coding conventions was used in developing the software. The WDM library of utility subroutines is used by any program that stores, retrieves, and updates data in the file.

The records in a WDM file are organized by data sets. Each data set contains a section for attributes, a section for pointers to groups of data within the data set, and a section containing the data groups. The attribute section identifies and characterizes the data and can be used for searches and retrievals. Examples of groups in a data set are months of hourly rainfall or cross sections along a reach of stream.

Five categories of pointers or links are used in the WDM file: record, directory, data set, group, and schematic. The schematic pointers use data set numbers and group numbers for pointers while the other pointers use records or locations within a record. Although pointer systems add complexity to the file, they also provide greater flexibility, more efficient storage, and quicker retrievals. More important, however, the complexity of the pointer system is transparent to the users and application programmers since all reads, writes, and updates are done with a library of subroutines.

Attributes

Attributes are not defined by a specific position in the record but kept dynamic so that space is not wasted and any combination of attributes can be used. There are about 300 predefined attributes that can be used, and the list continues to grow. All attributes are assigned an index number and placed in a separate read-only message file. Both the index and the attribute value are placed on the first record of each data set. Attributes are listed and defined on the message file. Defaults, maximum, minimum, check lists, size, type, and units are also included for each attribute. New attributes are added as needed and do not require software changes. Attributes also can be computed values such as mean, standard deviation, maximum, minimum, and serial correlation coefficient for a time series. Other examples of attributes include latitude, longitude, elevation, state code, station number, station name, hydrologic unit code, parameter code, statistic code, time units code, start date, end date, read/write flag, agency code, state FIPS code, aquifer type, and soils index. The current file contains 255 attributes.

Data Set Types

All data in a WDM file are placed in one of the following data sets:

(1) File definition data set,
(2) Directory data set,
(3) Raster (grid) data set,
(4) Table data set,
(5) Schematic data set,
(6) Time-series data set, and
(7) Space-time data set.

The file definition data set is the first record of a WDM file and is always required. At least one directory data set is needed to provide pointers to any of the optional types of data sets. Directory data sets are added dynamically as needed. The remaining types of data sets are optional and contain the basic data. The configuration of the data sets is illustrated in figure 1. The arrows indicate which type data sets can be used to reference other type data sets.

File Definition Data Sets

For the file definition data set, the first value represents the version of the software used to create the file and is used to check that the file is a WDM file. This number will be incremented by one for later WDM designs that are not compatible. Software will automatically update the file for upward compatibility but downward compatibility will not be maintained.

A date from the operating system's utilities and a name from the user are put on the file definition data set for documentation and reference. Since the file can be dynamically expanded, the current file size in records is stored. The next free record also is stored and will be one greater than the file size unless some data sets have been deleted. When data sets are deleted, the next free record points to the first record of the deleted data set. Also stored in the first record

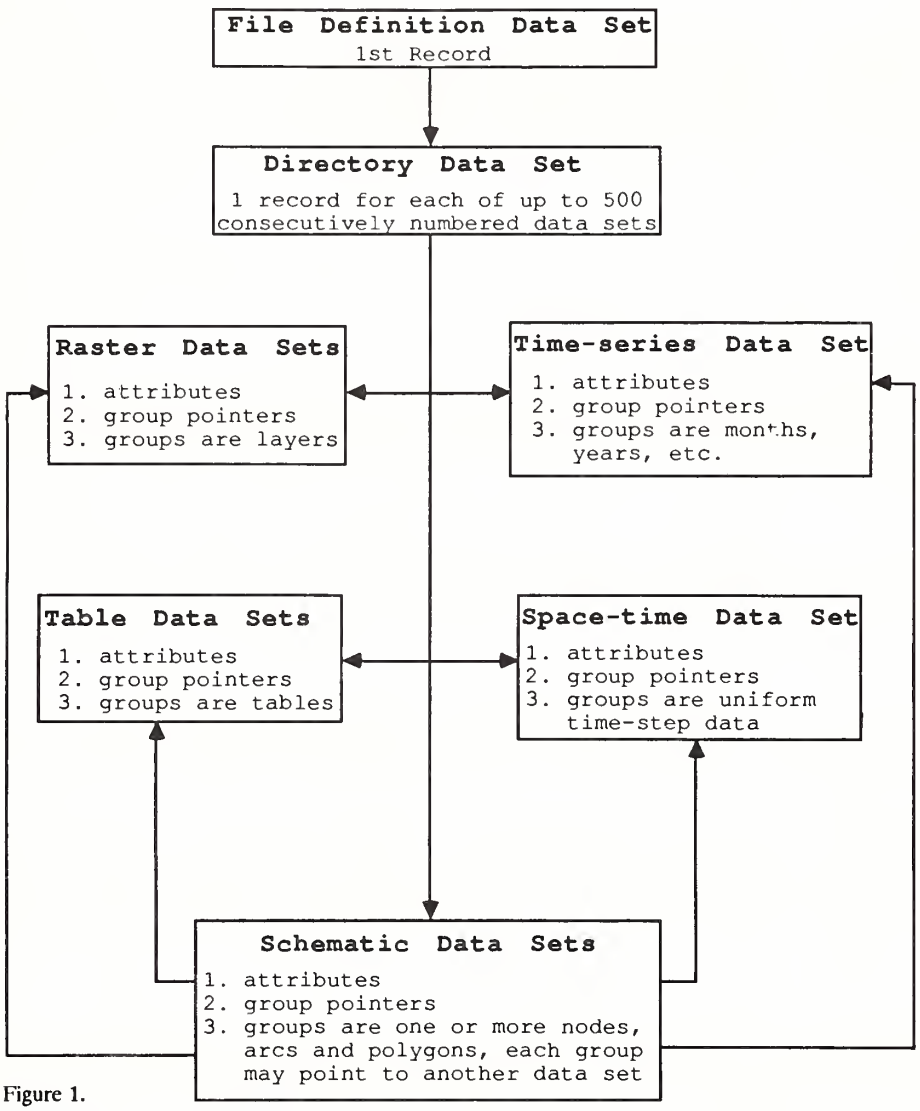


Figure 1.
WDM file configuration scheme.

are counts for the number of each type of data set and the most recently added data set number for each type of data set. All data sets of the specific types can be found using the first or last data set number of a specific data set type and the forward or backward primary pointers that exist for each data set.

The last string of numbers on the file definition data set are pointers to the directory data sets. Each directory data set contains 500 pointers to the individual optional data sets. Thus, if a file contained 64 directory data sets and each directory data set was full, 32,000 data sets could be referenced, which exceeds the capacity of most disks. A WDM file could fill an entire 600-megabyte disk, or it could be a small 10-record file containing two time-series data sets each with 5 years of daily streamflow.

Directory Data Sets

Up to 500 record pointers to any optional data set are found on the directory data sets. The first directory data set is for data sets 1 to 500, the second directory data set is for data sets 501 to 1000, etc. To illustrate the pointer system, assume data set number 2760 is requested by the software. Then, 2760/500 equals 5 with 260 as a remainder. In the file definition data set, the fifth item in the directory pointers contains the record number of the appropriate directory data set. The 260th item in the directory data set contains the record number of the first record for the requested data set.

Schematic Data Sets

The purpose of the schematic data sets is to reference and link polygons, arcs, nodes, and other data sets. Arcs are strings of x-z coordinates that could define a location of a road, center line of a stream, political boundary, etc. Often boundaries are shared, such as the divide between two drainage basins or a political boundary and the center of a stream. A polygon is an arc or a set of arcs that close, such as a drainage basin boundary. Nodes are points. The type of coordinates used are defined by attributes. The coordinates can be real or integer. Integer values in combination with appropriate transformations may be used to maintain greater precision. The coordinates will be stored in the DLG format (NCIC 1985). The details of this data set are currently being designed.

Table Data Sets

Table data sets store one or more, related or unrelated, tables of a format described on the message file used by ANNIE or a special WDM data set. Tables have header information, the order, type, and space described on the message file. Tables can have up to 16 fields of any mix of integer, real, double precision, or character data. Character fields must be in increments of words (4 bytes). A restriction to 16 fields is currently needed for the software to place the table on the display terminal for processing. The maximum number of rows is more system-dependent and based on the buffer for table input and output. The order, type, and space for data in the table also are stored on the message file.

Table data sets are not useful until specific application programs have been written, although the only requirement is the table description on the message file. A special interactive program is used to create the specification on the message file. Once created, a single subroutine call allows the user to interactively edit and plot data in the table.

Each table in a data set is a group and has a group pointer based on a table number. Boolean searches similar to attributes for data sets cannot be done with table header information unless specially set up by an applications programmer. If such searches are needed, one table per data set could be used with the table header information entered as data set attributes.

Tables have been programmed for channel cross sections, hydraulic properties of a cross section, and HSPF model parameters.

Continuous Time-Series Data Sets

There are several types and forms of time-series data. Most data have specific dates when they are measured or calculated. Some time series are synthesized and are independent of a specific date, such as design storm hydrographs. Some models use data in two or more time steps, such as a nonpoint loading model that may simulate on a daily time step until a major storm occurs and then it may simulate on a 15-minute time step. Some data are measured or calculated on a uniform time step, while other data are collected at random intervals. Water-quality samples are often collected at monthly, quarterly, or random intervals. With microprocessors at data-collection sites, data are being collected whenever a change in the measured variable occurs. Some time-series data need flags to indicate whether it was measured or estimated data, how the data were estimated, or periods of missing record.

To store all the various types of time series, yet keep the file format simple and efficient, each group is stored in blocks with each block preceded with a block control word. The block control word is a 32-bit integer with a bit pattern to represent five variables:

Variable	Value Range	Pattern Length
Number of values	0-32767	15 bits
Time step for the block	0-63	6 bits
Time-step units code	1-6	3 bits
Compression code	0-3	2 bits
Quality of data code	0-31	5 bits

Time unit codes range from 1 to 6 for seconds, minutes, hours, days, months, and years. The compression code is nonzero if the data in the block is compressed by one of three schemes: all values the same; range linearly between first and last value; or range nonlinearly with the type of nonlinear function defined in the data block. When the quality of the data is important, the quality-of-data code is set to a nonzero value. Quality codes identify missing data, several forms of estimated data, or data with missing time distributions.

In addition to the time-series data being put into blocks, blocks are put into groups for the pointer system within the data set. Groups would be for each hour, day, month, year, or century, as specified by one of the attributes. Another attribute is the beginning time for the first group of data. An attribute is also used to indicate whether the data is an average or cumulative over a time step or an instantaneous value.

As with all the other records, primary and secondary, forward and backward pointers are the first four positions. The primary pointers are used to link all the time-series data sets. These pointers contain data-set numbers of other time-series data sets. The primary forward pointers are used to reference the time-series data sets beginning with the first time-series data set number as defined in the file directory record. The primary backward pointer is used to re-link the data sets when a time-series data set is deleted. The secondary pointers are used to link the records within the data set since these are not necessarily contiguous records in the WDM file.

The format for a time-series data set is designed so that data, search attributes, and pointers can be added, updated, and deleted at any time and in any order. The result is a more complex format with most of the positions relative to pointers and size parameters. The complexity, however, is transparent to the applications programmer and the user since all reads, writes, and updates are done through a group of utility subroutines.

Raster (Grid) Data Sets

Raster or grid data sets can be used to store spatial data from a grid pattern. The data set assumes a uniform spacing between values. Any nonuniform spacing would have to be made by the receiving software but could use a table data set to indicate the spacing. The location and spacing of the grid is specified in the attributes.

The values in each grid can be real numbers, double precision numbers, integers, half-word integers, or single characters. If a multiplier is supplied as an attribute, integers can be returned as real numbers, the integer value times the multiplier. An attribute for an offset also can be used to add a positive or negative number to each grid value.

Space-Time Data Sets

Space-time data sets are similar to grid data sets except one dimension is time. Finite-difference and finite-element models use this data set when values for all nodes or elements are needed at a point in time. Time-series data sets could be used, but retrievals would be inefficient. Data are placed in groups of uniform time steps. Up to 200 time-step changes can be stored per data set, but the number of uniform time steps is only limited by the capacity of the computer and disk. The number of nodes is constant and cannot be greater than 10,000 per data set. Retrievals specify a time, starting node, and skipping factor.

SUMMARY

Managing data necessary for water-quality modeling is a very difficult problem. The cooperative efforts of several Federal agencies have produced a system that has begun to address the problem and has demonstrated that the approach has much promise. The system has been implemented with several water-quality models.

LITERATURE CITED

Baker, Claud H., Jr., and Donald G. Foulk. 1984. WATSTORE User's Guide, vol. 2, Ground-water file. Open-File Report 75-589, Revised April 1984, U.S. Geological Survey, Reston, VA.

Date, C.J. 1985. An Introduction to Database Systems, vol. I, Addison-Wesley Publishing Company, 639 pp.

Database Analyzer for Pesticide Evaluations (DBAPE). 1988. EPA Internal Report developed under contract submitted to Athens Environmental Research Laboratory, Office of Research and Development, Athens, GA.

Edwards, M.D. 1980. NAWDEX: A Key to Finding Water Data, U.S. Geological Survey, National Water Data Exchange pamphlet, 15 pp.

Gleason, M.N., R.E. Gosselin, H.G. Hodge, and R.P. Smith, eds. 1969. Clinical Toxicology of Commercial Pollutants. Williams and Wilkins, Baltimore.

Hartley, D., and H. Kidd. 1983. Agrochemicals Handbook. Royal Society of Chemistry. Nottingham, England.

Hatch, Warren L. 1983. Selective Guide to Climate Data Sources, Key to Meteorological Records Documentation No. 4.11, National Climatic Data Center, Asheville, NC, 338 pp.

Hydrologic Engineering Center, U.S. Army Corps of Engineers. 1983. HECDSS User's Guide and Utility Program Manuals, Davis, CA.

Lumb, Alan M., and John L. Kittle, Jr. 1985. ANNIE - Interactive Processing of Data Bases for Hydrologic Models, Proceedings of the International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology, Los Angeles, California.

Nash, R.G. 1988. In Herbicides in Soil. R. Grover, ed., vol I, pp. 131-169. CRC Press, Inc., Boca Raton, FL.

National Cartographic Information Center (NCIC). 1985. Digital Cartographic and Geographic Data. U.S. Geological Survey pamphlet, Reston, VA, 4 pp.

Smith, Anne Q., and C. Robert Clauer. 1986. A Versatile Source-Independent System for Digital Data Management, EOS, vol. 67, No. 15, pp. 188-189.

Treinish, Lloyd A. 1984. A General Scientific Information System to Support the Study of Climate-Related Data, NASA/GSFC TM 86152.

USDA Soil Conservation Service. 1982. National Resources Inventory Database. USDA Soil Conservation Service, Ames, Iowa.

USDA Soil Conservation Service. 1985. User Manual Interactive Soils Databases: National Soil Survey Area Database, Soil Interpretations Record Database and Plant Name Database. USDA Soil Conservation Service, Fort Collins, CO.

Windholz, M., S. Budavari, R.F. Blumetti, and E.S. Otterbein. 1983. The Merck Index. Merck and Co., Inc., Rathway, NJ.

DISCUSSION OF THE PAPERS PRESENTED IN TECHNICAL SESSION 8, PART 3: DATA BASE DEVELOPMENT

Daniel Hoggan[1], Presiding
Donna Falkenborg[2], Recorder

PAPERS DISCUSSED

Use of Soil Survey and Other Data Bases in the Modeling of Leaching from Agricultural Sources by A. Breeuwsma and J. Bouma

Data Management for Water-Quality Modeling Development and Use by A.M. Lumb, R.F. Carsel and J.L. Kittle, Jr.

SPECIFIC QUESTIONS AND COMMENTS

Question: (D. Jackson, Susquehanna River Basin Commission, Harrisburg, Pennsylvania) What is the status of components of various kinds of data going into the GIS system?

Response: (A. Lumb, USGS, Reston, Virginia) Data on various soils properties is digitized, but digitizing boundaries of all those soils is a slow process and not a lot has been done. Vegetal cover data is from remote sensing; tapes are purchased. Occasionally land cover data has been digitized by state agencies but not in a particularly consistent fashion.

Question: (E. Casman, Interstate Commission on the Potomac River, Baltimore, Maryland) When you're dealing with extracting parameters from a single cell that had different land uses and soil types, did you aerially weight the parameters or did you use some other system to combine properties?

Response: (A. Breeuwsma, Soil Survey Institute, Netherlands) We didn't weight them, we combined the data for each pixel. Thus you have data in the data base for each pixel; the soil type and also the land use.

Question: (E. Casman) Is there no need to average?

Response: (A. Breeuwsma) No, but that was an important point. Averaging is often deadly when talking about leaching problems, so we need to have point data as much as possible. However in the example I show we used average values. In the future we will be able to take care of the distribution of parameters within soil type and also within the grid as far as the input of pollutants is concerned.

Question: (Audience) What kind of time step was the model running on?

Response: (A. Breeuwsma) For the phosphate study we used a very simple model. We just divided the phosphate input by the phosphate options capacity. So the time step in this case was irrelevant. It was not a physical or hydrological model in this case. We just showed we had a certain average annual precipitation. This was used in order to make calculations as to the

[1]Daniel Hoggan, Professor, Utah Water Research Laboratory, Utah State University, Logan, Utah.
[2]Donna Falkenborg, Logan, Utah.

amount of phosphate dissolved from the manure or animal waste that has been added. It is a different situation than the models that have been discussed.

Question: (D. DeCoursey, USDA-ARS, Hydro-Ecosystems Research Unit, Fort Collins, Colorado) Is anyone in your country looking into providing a description of the soils that will enable us to incorporate some of the concepts of macroporosity such as was described yesterday? Is anyone looking into characterizing the soils?

Response: (A. Breeuwsma) Yes, that's a big effort now being made. Henricks has recently studied the regular flow patterns in sandy soils and it's very surprising that even in very homogeneous, sandy soils infiltration rates are different. This has something to do with the water repellency and what they are trying to do now is characterize the input variation, infiltration rate, and relate it to different soil types. This is going to be a very important issue, especially when pesticide leaching is the target. Due to irregular flow, pesticide moves much faster, even in sandy soils, than we would expect.

Poster Presentations

A MODEL FOR SIMULATION OF NITROGEN DYNAMICS IN SOIL AND NITRATE LEACHING

Lars Bergström[1], Per-Erik Jansson[2], Holger Johnsson[3]

ABSTRACT

A soil nitrogen model emphasizing mineral-N dynamics and nitrate leaching is presented; both are shown to be described satisfactorily with the model. Driving variables are generated with a water and heat model using standard meteorological data. Site-specific input data to the nitrogen model should preferably be estimated from standardized measurements, while it should be possible to obtain basic information about cropping systems without measurements at the investigated site. The model can be run on IBM-PC, VAX or PRIME computers.

INTRODUCTION

Nitrate leaching from arable land has long been identified as a serious environmental problem. In order to develop efficient and useful counter measures, a better quantitative knowledge is needed of how climate, soil types, and cropping systems interact to influence the water and nitrogen balances in soil. Mathematical simulation models are excellent tools for studying these complex problems. By using models we are able to separate the influence of different processes and their regulating factors on the soil nitrogen status and thereby interpret results from field studies where many factors often covariate.

The model described (Johnsson et al. 1987) was developed within the project "Ecology of Arable Land". The model emphasizes mineral-N dynamics and nitrate leaching in agricultural soils. The focus of model development was to facilitate numerous applications, with a model resolution compatible with information generally available from agricultural field research.

MODEL DESCRIPTION

All major N-flows occurring in agricultural soils are considered in the model (fig. 1). A water and heat model (Jansson and Halldin 1979) describing water and temperature conditions in soil, provides driving variables for the model. The soil profile is divided into layers which are considered as biologically and physically homogeneous.

The N-flow rates in soil are regulated by the temperatures, water contents and water flows that are simulated by the water and heat model. Thus necessary information to run the model could be limited to climatic variables, soil properties, plant N-uptake and inputs of nitrogen fertilizers.

The mineral-N pool consists of ammonium and nitrate. Organic-N is classified as litter (e.g. undecomposed above- and below-ground crop residues and microbial biomass), manure-derived faeces and humus. Nitrogen can be mineralized from or immobilized to litter and faeces depending on the C/N ratio. Nitrogen is always mineralized from humus due to its low and stable C/N ratio. The plant component includes nitrogen in above- and below-ground biomass. Roots and crop

[1]Lars Bergström, Research Leader, Swedish Univ. Agric. Sci., Uppsala, Sweden.

[2]Per-Erik Jansson, Professor, Swedish Univ. Agric. Sci., Uppsala Sweden.

[3]Holger Johnsson, Research Assistant, Swedish Univ. Agric. Sci., Uppsala, Sweden.

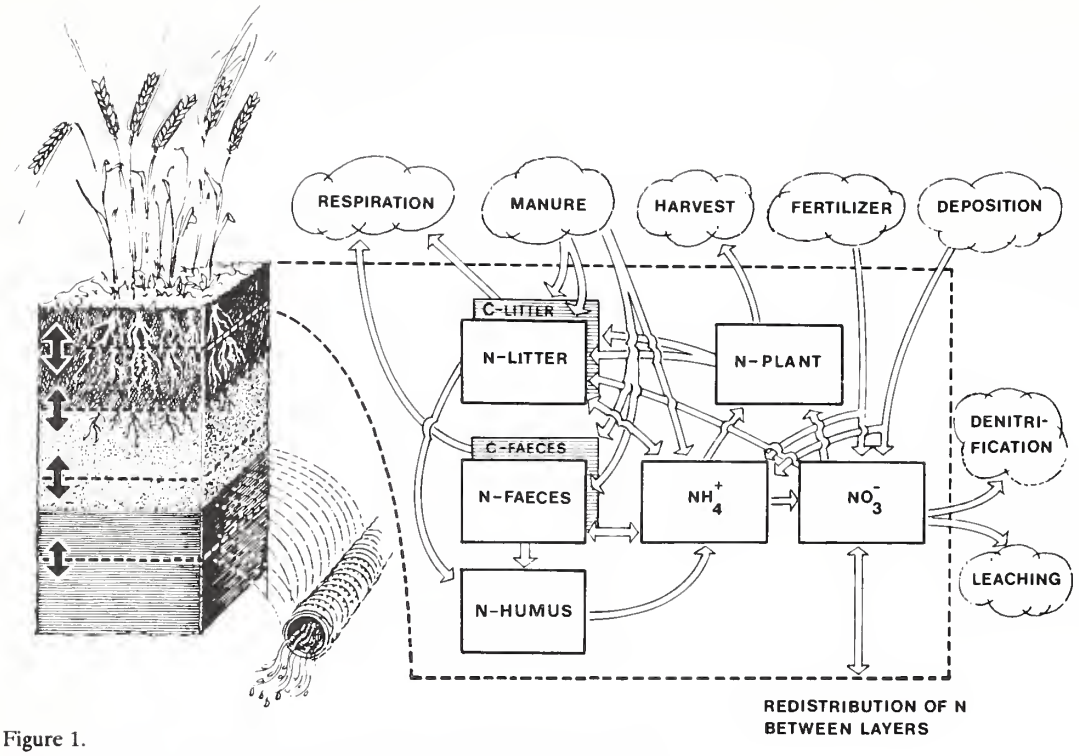


Figure 1.
Model structure. Parts within the broken line represent the surface layer. Subsurface layers show the same structure, but no direct transfers from fertilizer and deposition (from Johnsson et al. 1987).

residues are assigned certain C/N ratios which characterize their nitrogen contents before incorporation in soil.

Manure, inorganic fertilizer and atmospheric deposition are external inputs to the uppermost soil layer. Leaching, harvest, and denitrification constitute nitrogen losses from the soil. Nitrate in solution can be transported between soil layers, to drainage tiles or to deeper groundwater.

MODEL APPLICATIONS

In a first model test it was important to choose a field where information on all major parts of the model was available, enabling a thorough test of the model. The experimental site for the project "Ecology of Arable Land" was thus suitable since measurements of most nitrogen transports and transformations occurring in agricultural soils were performed. Barley, with and without addition of N fertilizer, was chosen as test crop (Johnsson et al. 1987). In a later application, data from a perennial grass ley system was also used (Bergström and Johnsson 1988). Comparison between simulated and measured nitrate leaching and mineral-N dynamics was done for periods of between three and four years. Nitrate leaching was measured from 0.36-ha tile- drained plots. The mineral-N content of the soil was determined down to a depth of 1 meter.

The model has been tested with experimental data from several other sites in Sweden covering large areal scales and long periods (Jansson and Andersson 1988, Gustafson 1988), smaller fields with measurements of soil mineral nitrogen (Jansson et al. 1987) and fields where detailed information on both mineral-N dynamics and nitrate leaching was available (Borg et al. 1988).

RESULTS AND DISCUSSION

Experiences From The First Model Test

Results from the first application with barley showed a reasonable agreement between simulation and measurements. This applied to both mineral-N dynamics in the soil and nitrate leaching (fig. 2).

Both biotic and abiotic factors were shown to be important for the nitrogen dynamics: biotic factors mainly in the topsoil and abiotic factors in the subsoil. A better agreement between simulated and measured nitrogen dynamics was generally obtained for surface horizons than for deeper layers in the soil.

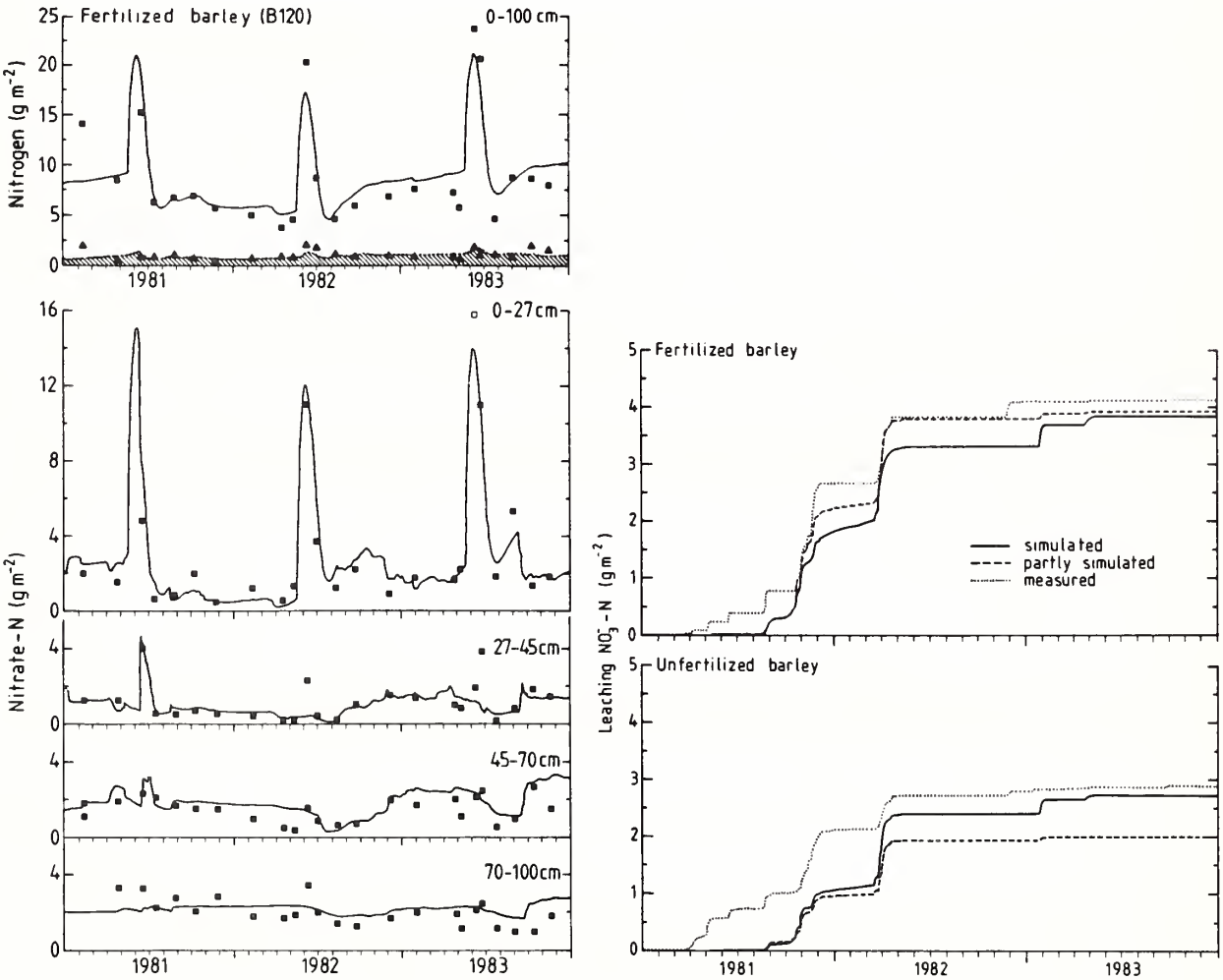


Figure 2.
Examples of model output. To the left: Simulated and measured storage of mineral-N in the soil at Kjettslinge. The crop was barley fertilized with 12 g-N/m^2. In the upper figure, squares represent total mineral-N content and triangles ammonium- N content. To the right: Simulated, partly simulated (measured nitrate concentration multiplied by simulated water flows from drainage tiles) and measured leaching from N-fertilized (12 g- N/m^2) and unfertilized barley at Kjettslinge (from Johnsson et al. 1987).

The temporal distribution of nitrate leaching indicated that both water-flow rates and flow paths were of importance. Simulated and measured depth distribution of nitrate in soil showed that a division of the soil profile into different layers is necessary to explain the coupling between soil mineral-N contents and nitrate leaching.

Future Work

Since mineralization and N-uptake by plants are normally the largest N flows in agricultural soils, it is of the utmost importance that these flows are properly calculated.

Estimation of N-mineralization in different soil types is still uncertain. The simulations showed the importance of knowing the amounts and quality of fresh easily decomposable organic matter in soil. Root turnover is one example of such a process which although it represents a substantial part of the N-supply of crops is rarely investigated in sufficient detail.

In the simulations reported here, N-uptake by the crop was modified until a reasonable agreement with measurements from the field was achieved. However, for future simulations, the model has to be improved so that predictions can be done without relying on detailed biomass estimations. It would be advantageous if N- uptake by the crop could be related to time-varying environmental factors such as: air temperature, global radiation and soil water supply. The present version of the model only considers nitrogen as the limiting factor for crop growth. We believe that there are good possibilities to improve the formulation of crop N-uptake since there are many results available from field studies describing variation in harvested-N with time.

To improve simulation of leaching it is necessary to thoroughly investigate how different soil properties, such as texture, formation of aggregates and cracks, influence percolation to drainage tiles and groundwater.

The model is developed so that as many properties and parameters as possible can be estimated from general knowledge, which is independent of each application. Future work will focus on trying to estimate necessary site-specific input data to the model from standardized measurements. For example, it should be possible to estimate both necessary physical properties of soils and mineralization capacities of soils, from standardized soil analyses. It would also be advantageous to work with crop scientists to develop information about important crop properties into readily-available data bases.

This would enable model applications for areas which are lacking detailed information, and would thereby significantly increase the possibilities for practical use of the model for fertilizer fate and leaching risk assessments. We believe that the model in its present form can be used successfully for practical purposes. It has already been shown to function well as a tool for evaluation of agricultural field research. In particular it has been used to synthesize knowledge gained in leaching studies and plant nutrition experiments.

LITERATURE CITED

Bergström, Lars, and Holger Johnsson. 1988. Simulated nitrogen dynamics and nitrate leaching in a perennial grass ley. Plant Soil 105:273-281.

Borg, Gunnar Ch., Per-Erik Jansson, and Börje Lindén. Simulation of nitrogen conditions in a manured and fertilized soil. (Submitted to Plant Soil)

Gustafson, Arne. 1988. Simulation of nitrate leaching from arable land in southern Sweden. Acta Agric. Scand. 38:13-23.

Jansson, Per-Erik, and Sven Halldin. 1979. Model for the annual water and energy flow in a layered soil. In Sven Halldin ed. Comparison of Forest and Energy Exchange Models, pp 145-163. Society for Ecological Modelling, Copenhagen.

Jansson, Per-Erik, and Rune Andersson. 1988. Simulation of runoff and nitrate leaching from an agricultural district in Sweden. J. Hydrol. 99:33-47.

Jansson, Per-Erik, Gunnar Ch. Borg, Lars-Christer Lundin, and Börje Lindén. 1987. Simulation of soil nitrogen storage and leaching. Applications to different Swedish agricultural systems. Swedish National Environment Protection Board Rep 3356, 63 pp.

Johnsson, Holger, Lars Bergström, Per-Erik Jansson, and Keith Paustian. 1987. Simulated nitrogen dynamics and losses in a layered agricultural soil. Agric. Ecosystems Environ. 18:333-356.

THE SOIL-CREAMS MODEL TO SIMULATE SOIL AND CHEMICAL LOSSES FROM AGRICULTURAL AREAS.

Peter F. Botterweg[1]

ABSTRACT

The SOIL-CREAMS model is a combination of two existing models developed respectively in Sweden and the USA. Our first experiences with the U.S. Department of Agriculture CREAMS model showed that, especially the hydrology part, was not adequate for the winter and spring situation in Norway. We therefore decided to exchange the hydrology module in the CREAMS model (Knisel, 1980) with the SOIL model (Jansson et al., 1980, 1988). The SOIL model has good routines for soil frost and snowmelt. This is important for we recognized in the field, that as much as 75% of yearly runoff may arise from snowmelt.

Three programs had to be developed to combine the two models because of differences in variable definitions, input/output formats, and units.

THE SOIL MODEL

The SOIL model was primarily designed to predict annual soil climate for biological purpose within the Swedish Coniferous Forest Project. The model is physically based and general enough to predict annual water and heat dynamics, including frost, for a variety of layered, unsaturated soils and vegetation covers.

Soil boundary conditions are supplied by submodels of snow dynamics, precipitation interception, evapotranspiration, root water uptake and net horizontal groundwater flow. Given a measured soil surface temperature, the model can simulate soil temperature and heat flow variations within the day. For long-term simulations driving variables include daily means or sums of air temperature, precipitation, relative humidity and windspeed. Evapotranspiration is predicted from input of either net or global radiation, or cloudiness, or duration of bright sunshine. Soil properties are determined by independent methods but most surface characteristics must be determined by optimization. Mass and heat balance schemes for the model are illustrated in figure 1.

The two partial differential equations for heat and water flow are solved with an explicit forward differencing method (Euler integration). This requires the soil profile to be approximated with a discrete number of internally homogenous layers. Slowly changing state variables are updated at larger time steps, and integration time step varies dynamically as a function of conditions (including frost occurrence) during simulation to minimize execution times. Water flow rates into the two top soil layers are used as tests.

THE CREAMS MODEL

CREAMS is a field scale model for Chemicals, Runoff, and Erosion from Agricultural Management Systems. Today it is probably the most frequently used model for estimating nonpoint source pollution. The model consists of three separate modules which treat respectively hydrology, erosion and losses of nutrients (P and N) and pesticides. For this study, the hydrology

[1]Peter F. Botterweg, Research Officer, Institute for Georesources and Pollution Research, P. O. Box 9, N-1432 Ås-NLHG, NORWAY

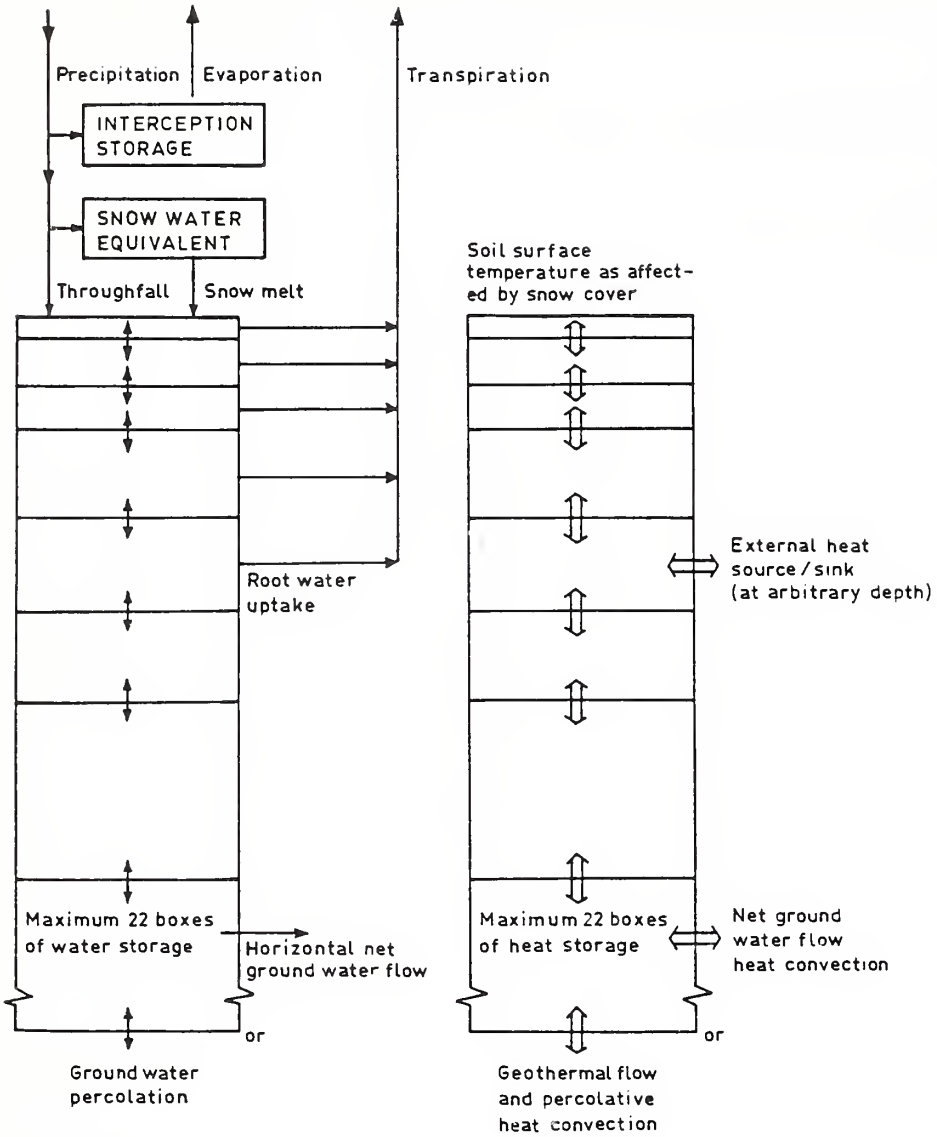


Figure 1.
Mass balance and heat balance of the SOIL model.

module has been replaced by the SOIL-model. The erosion component of the model is a modification of the USLE, an empirical formula derived from plot data that employs physically related factors. Losses of chemicals are calculated as a sum of losses by runoff, sediment transport, and leaching. The CREAMS model is expected to be well known and is not discussed in detail here.

The original CREAMS model gives information on sediment only in the report file. We wanted to have the possibility to use these data for other analyses and therefore made some changes in the program. Statements were added in the STROUT subroutine to write a "pass file" with the following parameters:

SDATE	(date of the storm)
K	(soil particle type)
FRACSMD	(fraction of particle K in sediment)
CONC	(concentration of particle K in runoff)

SEDCLY	(clay fraction in sediment)
SEDSLT	(silt fraction in sediment)
SEDSND	(sand fraction in sediment)
SEDORG	(fraction organic matter in sediment).

RUNNING THE MODEL

The SOIL-CREAMS model was run on an IBM-PC in batch mode. An interactive, user-friendly system for running the model has not been developed yet. The SOILCRM program reads the run number from the SOIL model, defines file names and run files for the programs that follow. A flow chart for a complete simulation is given in figure 2.

The soil parameters for the model run are read from the file SOILP.DAT. This file can be created by the PFFIT program from a soil data bank that contains pF-values, texture, % organic matter, and saturated conductivity.

Driving variables can either be estimated by analytical functions or be read from an external file. With help of switches and parameter values a wide range of different combinations of driving variables can be used.

The parameter file for the SOIL program is extensive, with 153 parameters (but not all of these are necessary for each run). The setting of 24 switches gives the model great flexibility; it can easily be adjusted to specialized research.

The summary report file includes parameter values, switch settings, and minimum ,maximum and cumulative values of the user-selected output variables. This report file also gives soil parameters describing each soil layer. The output from the CREAMS erosion and chemistry modules are summarized by month and written to the soil report file. All output data are stored in one SASDATA SET.

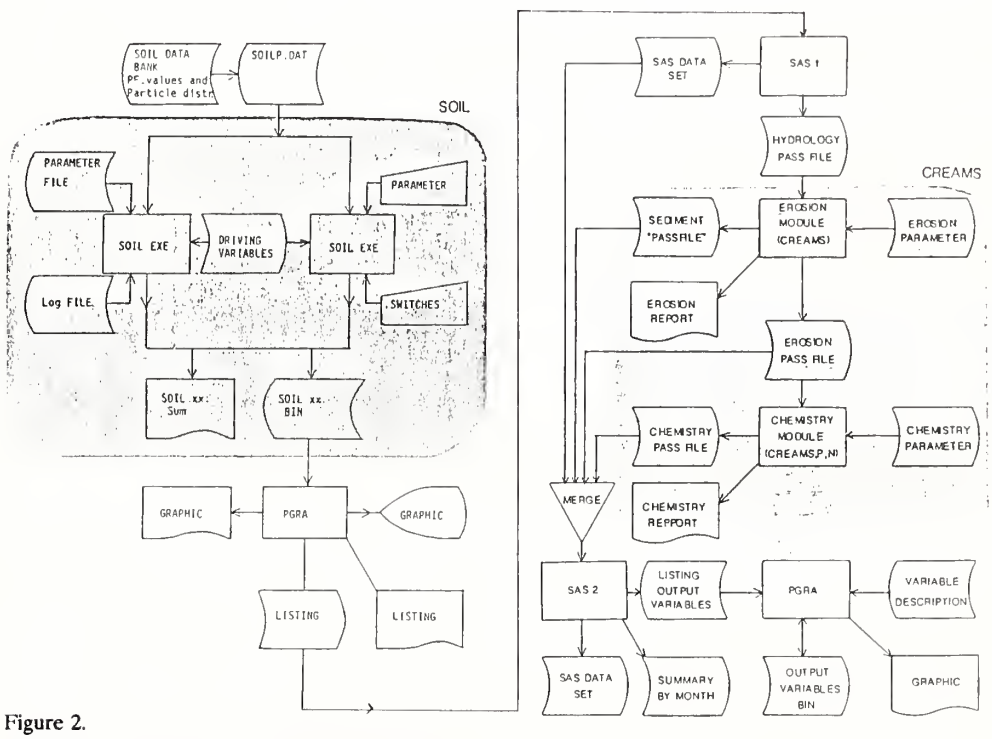


Figure 2.
Flow chart for the SOIL-CREAMS model system.

The SOIL-CREAMS model has been preliminarily tested on a heavy clay soil. The hydrology part has been successful with reasonable simulation of surface runoff and drainage compared to measured data. The simulation of erosion was not as good as expected. It has been necessary to change the American values for the USLE factors quite a lot to reduce the soil loss level by a factor of ten, to the level of measured values. The distribution of total soil loss over the year was satisfactory, but the snowmelt period needed special attention.

Our approach, combining two different models has shown promising results, and we will continue this work.

LITERATURE

Jansson, P., and S. Halldin. 1980. Soil water and heat model: Technical description. Swedish Coniferous Forest Project, Technical Report 26.

Jansson, P. 1988. Manual and help libraries for the SOIL model. Department of Soil Sciences, Swedish University of Agricultural Sciences, Uppsala.

Knisel, W.G. 1980. CREAMS. A field scale model for chemicals, runoff, and erosion from agricultural management systems. U.S. Department of Agriculture, Conservation Research Report No. 26, 640 p.

RZWQM - A Model for Simulating the Movement of Water and Solutes in the Root Zone

Donn G. DeCoursey[1] and Kenneth W. Rojas[2]

ABSTRACT

This paper describes an unsaturated zone water quality model being developed by the USDA-Agricultural Research Service. The model simulates the physical, chemical and biological processes responsible for the movement of water, nutrients, and pesticides over and through the root zone of a point in an agricultural field. Brief descriptions of each of the six major processes are described; representative output illustrate important features.

INTRODUCTION

Agricultural development and the intensive use of fertilizers and pesticides have led to contamination of ground water supplies. Development of improved management strategies to mitigate these effects of agricultural development will require use of mathematical models that simulate the physical processes responsible for the pollution. The root zone water quality model (RZWQM) being developed by USDA-Agricultural Research Service (ARS) is one such model.

The model is being developed as an aid to field research in studying alternatives to conventional agricultural practice that could reduce the potential for ground water pollution. It consists of three major subsystems: an input file generator, the model of physical, chemical and biological processes, and an output report generator. The conceptual framework and some of the code were provided by a group of ARS scientists. The processes and their authors are: physical processes (Laj Ahuja, Alan Hjelmfelt, Charles Hebson, Ken Rojas), nutrient processes (Marvin Shaffer, Charles Hebson), soil chemistry (Marvin Shaffer, Ken Rojas), pesticide processes (Ralph Nash, Don Wauchope, Guye Willis, Les McDowell), plant growth (Jon Hanson, Allan Jones, Ken Rojas), and management (Jim Schepers, Ken Rojas).

The following material describes each of the six major processes and input and output subsystems. Representative output illustrate important features.

THE PHYSICAL, CHEMICAL AND BIOLOGICAL PROCESSES

The central computational feature of RZWQM consists of the six major processes mentioned above. These processes are fully integrated into RZWQM to provide the necessary feedback to describe the interaction that takes place as water and solutes move in response to the variety of driving forces and constraints imposed on the flow system by nature and agricultural management.

Physical Processes

The physical processes are the internal hydraulics and hydrologic processes that interact to simulate the movement of water and solutes. The major subprocesses include: infiltration; chemical transport during infiltration; transfer of chemicals to runoff during rainfall; water and

[1]Donn G. Decoursey, Research Leader, USDA-ARS, Hydro-Ecosystems Research Unit, Fort Collins, CO 80522
[2]Kenneth W. Rojas, Research Associate, Colorado State University, Fort Collins, CO 80523

chemical flow through macropore channels and their absorption by the soil matrix; soil hydraulic properties estimation from bulk density and 1/3, or 1/10 bar water content; diffusive and advective heat flow; potential evaporation from the soil and residue surfaces; potential transpiration; root water uptake and soil water redistribution; and chemical transport during redistribution. Infiltration and macropore flow during precipitation events are simulated by layered and radial Green and Ampt expressions respectively. Moisture redistribution between events is simulated using the Richards' equation for unsaturated flow. Figures 1A and 1B are similar three dimensional plots showing simulated soil water content of a Bethany loam soil in Oklahoma as a function of time and depth in the soil profile. Note the non linear depth scale (all plots are similar in this respect). Input weather data for these examples (precipitation amount and duration; solar energy; maximum, minimum, and dewpoint temperatures; and wind run) were generated using a climate generator (Nicks and Lane, 1989). Figure 1A assumes no macropore soil features. Figure 1B assumes the soil to have a macropore soil structure. Macropore flow was responsible for the spikes that appear on the second soil horizon about day 190 in Figure 1B. Macropore flow is not evident at other times of the year because the soil profile is nearly saturated thus the macropore flow moves completely through the profile. It shows up as spikes on about day 190 because the soil was very dry at this time, due to root water uptake, thus some of the macropore flow was absorbed by the soil matrix raising the soil moisture content. In addition to root water uptake, the figures show surface drying due to evaporation.

Soil Chemistry

Soil chemistry processes consist of those necessary to describe the soil inorganic chemical environment in support of nutrient and pesticide processes. The inorganic processes include bicarbonate buffering; dissolution and precipitation of calcium carbonate, gypsum, and aluminum hydroxide; ion exchange involving bases and aluminum; and solution chemistry of aluminum hydroxide. The chemical state of the soil is characterized by soil pH and solution concentrations of aluminum and other cations depending upon pH. Collectively these processes consist of a set of non-linear equilibrium equations that are solved by a Newton-Raphson technique.

Nutrient Processes:

The nutrient processes define transformations of absorbed and soluble nutrients at all times. Given initial levels of organic matter, crop residue, and nutrient concentrations the submodel simulates the decomposition of soil organic matter and crop residues, the mineralization, nitrification, immobilization, denitrification, and volatilization of appropriate nitrogen and phosphorus species; and the absorption/desorption processes of both phosphorous and nitrogen. Levels of soluble nutrients are used in estimating crop growth, nutrient extraction in surface runoff, and movement below the root zone. Surface absorbed materials are subject to erosion. The organic matter/microbial population/nitrogen cycle is a complicated set of non-linear ordinary differential equations based on an Arrhenius formulation of interaction dynamics. It is solved by a standard fourth order Runge-Kutta algorithm. Again using the same example as in Figure 1, Figure 2 shows nitrate concentration in the soil water after application of 170 kg/ha on day 120. Figure 3 shows the concentration profile on three different days in the season with and without macropore flow. Note in Figure 3 the surface drying effect keeping, and in fact increasing, concentrations in the surface layer. Later it is completely leached below the surface layers. Also note in Figure 3 the effect of macropore flow on the location of peak concentration on day 300. Surface water moving through the macropores bypassed much of the soil matrix thus less water was available to move the nitrate and it remained higher in the soil profile.

Pesticide Processes

Pesticide processes consist of those necessary to estimate the transformation and degradation of pesticides (1) on the plant, crop residues, and soil surface and (2) in specific soil layers.

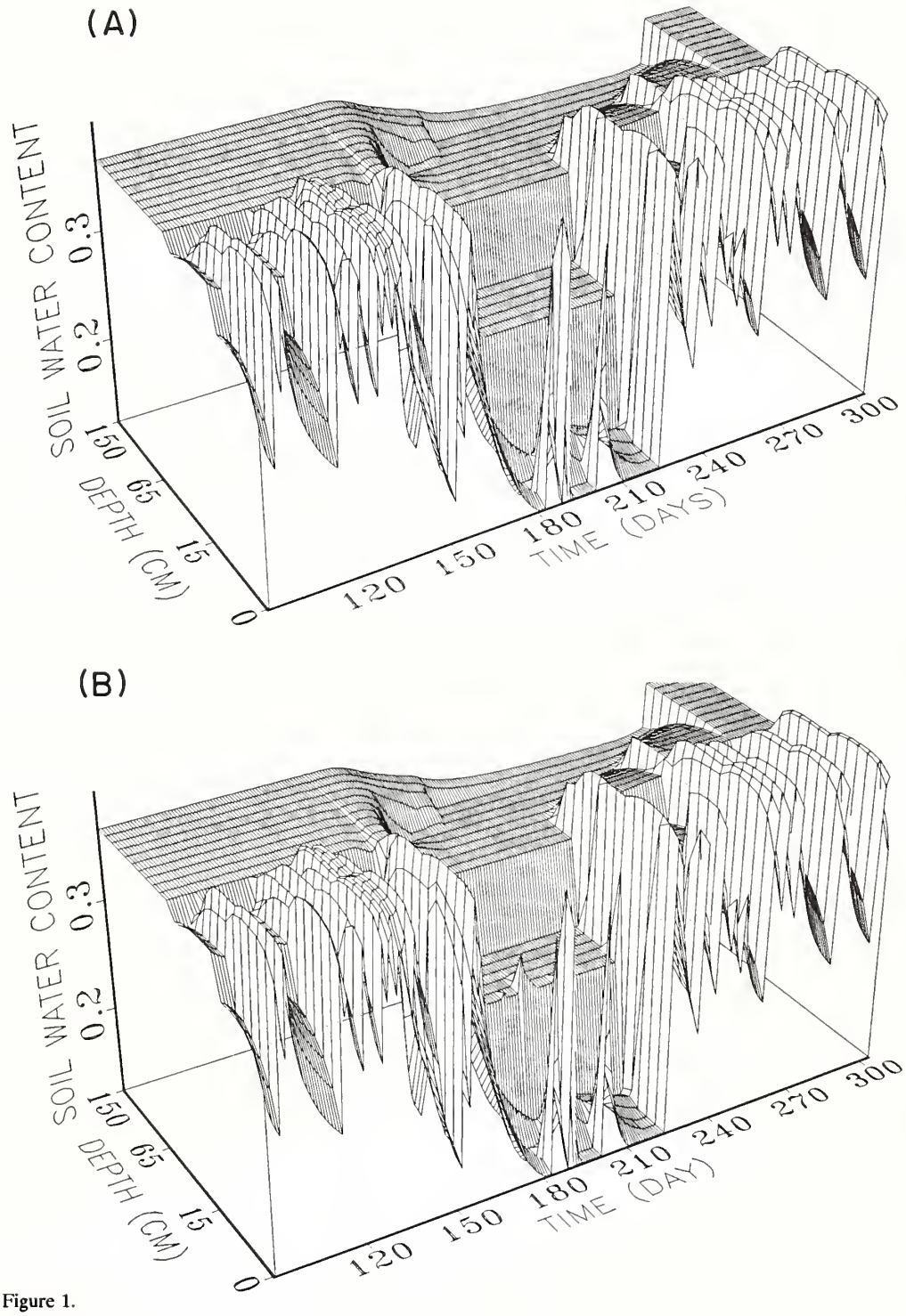


Figure 1.
Soil moisture content of a Bethany loam soil during the growing season showing the effects of macropore flow, surface evaporation, and soil water uptake by plant roots. (A - no macropore soil structure; B - assumes soil has a macropore soil structure).

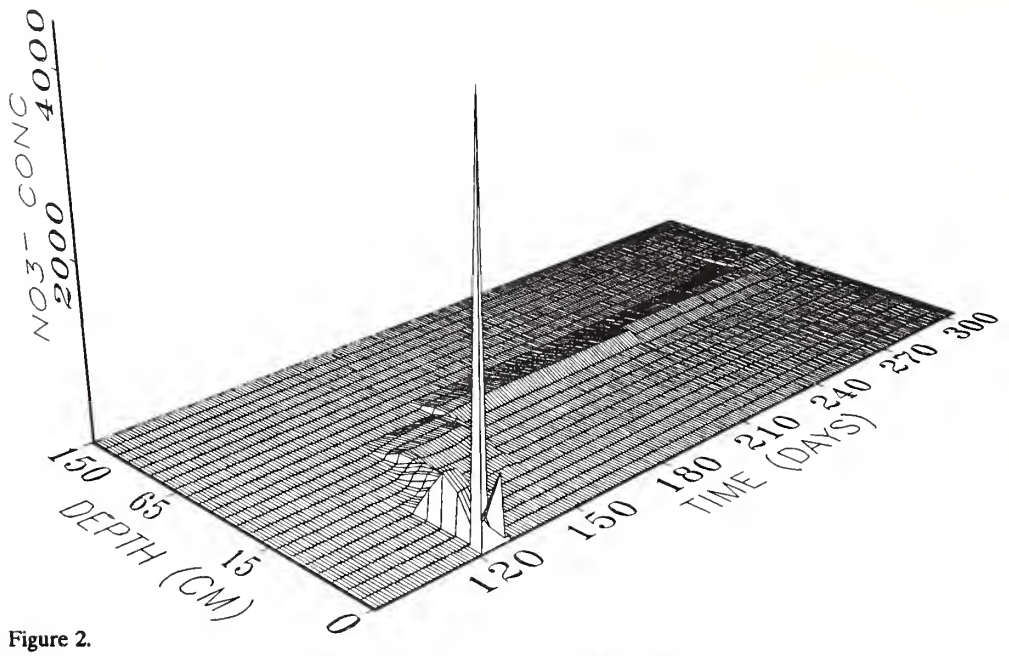
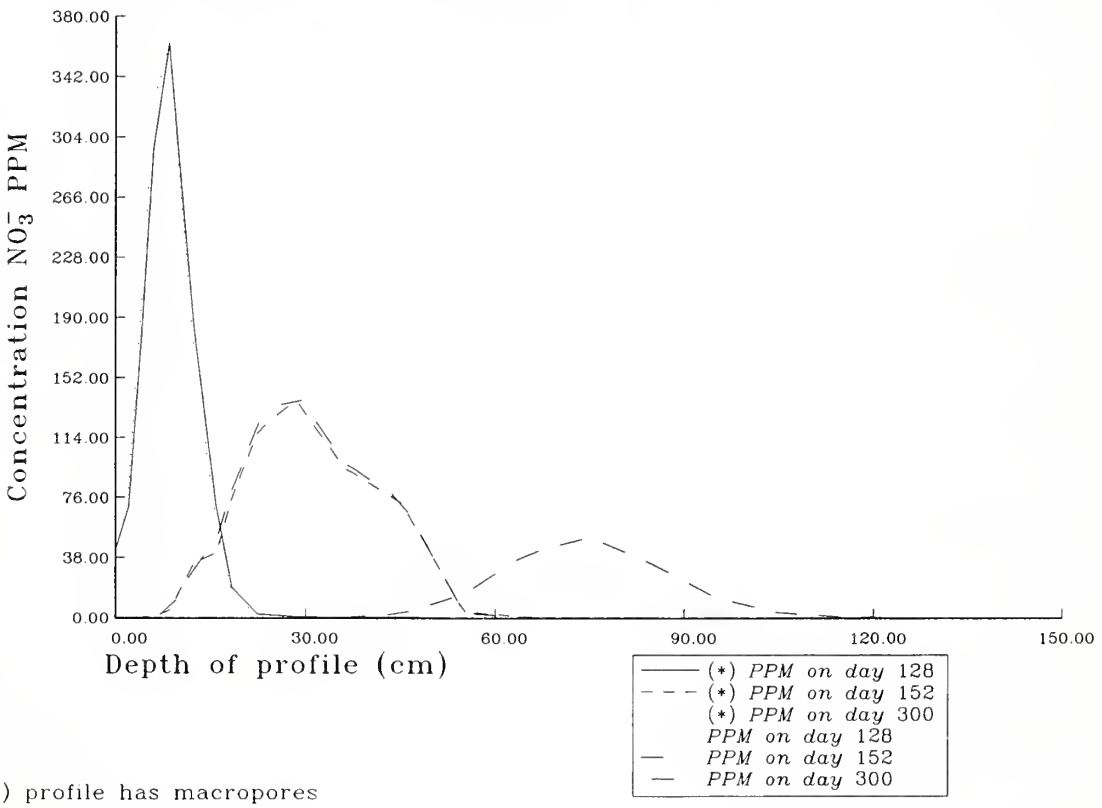


Figure 2.
Nitrate concentration in soil water after application of 170 kg/ha on day 120.



(*) profile has macropores

Figure 3.
Nitrate concentration profiles on days 128, 152, and 300 with and without macropore flow. 170 kg/ha of nitrate fertilizer was applied on day 120.

Depending upon the site; and given the plant, crop residue, soil and pesticide characteristics and environmental conditions; the model simulates the amount of pesticide reaching the soil surface

and the amounts absorbed and moving though each soil layer. Several dissipation methods are available: simple lumped dissipation, two compartment dissipation for highly volatile pesticides, individual degradation pathway dissipation, and daughter product formulation dissipation. Modifiers to further enhance the system are provided for all the above dissipation methods. The modifiers include pesticide formulation, plant leaf characteristics, residue interaction, and soil surface microtopology. Equilibrium and kinetic adsorption/desorption isotherms are used to determine the balance between adsorbed and solution phases. The structure of the submodel enables us to simulate the multi-site or mobile/immobile water concepts of adsorbed chemical transport (both concepts lead to the same mathematical expressions). Figure 4 shows the fate of 3.3 kg/ha of atrazine applied on day 126. The soil and climatic inputs are the same as described in Figure 1. Atrazine has a half-life of about 60 days and a K_{oc} of 100 μg/g. Note both the dissipation and leaching of the material. Compare the leaching to that of nitrate in figure 3 which shows nitrate (not adsorbed at all) considerably deeper in the soil profile. Even though atrazine is partially adsorbed, the model shows that 70 g/ha would have moved through macropores to depths exceeding 150 cm (see DeCoursey et. al., 1989). In contrast to atrazine, which is quite mobile, the fate of two other chemicals is shown in figures 5 and 6. Figure 5 shows concentrations of paraquat applied at a rate of 1.1 kg/ha in conjunction with the atrazine for weed control. It has both a very long half-life (500 days) and a very high K_{oc} (1 X 10^5). Most of the drop in concentration shown in Figure 5 is degradation. Figure 6 shows concentrations of chlorpyrifos applied on day 190 for insect control at a rate of 4.4 kg/ha. Chlorpyrifos has a high K_{oc} (6070) but a very short half-life (30 days). Difference in fate of these two chemicals as a result of half-life degradation rate is obvious. Both chemicals have high K_{oc}'s thus they remain at or near the surface.

Plant Growth Processes

The crop growth submodel describes crop growth to the extent that it estimates (1) specific yield of fruit, forage, or root crops; (2) soil cover conditions; (3) uptake of nutrients and pesticides as moved in transpiration uptake of water; (4) total dry mass of material grown (by plant parts) and its death or abscission; (5) the effects of water, nutrient, temperature, and pesticide stress; and

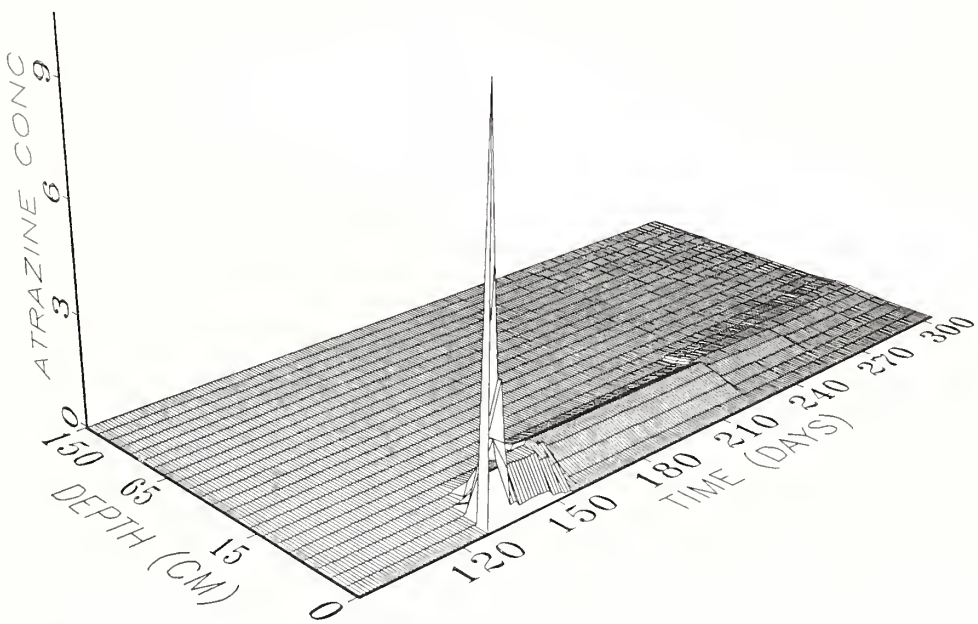


Figure 4.
Atrazine concentration in soil water after application of 3.3 kg/ha on day 126.

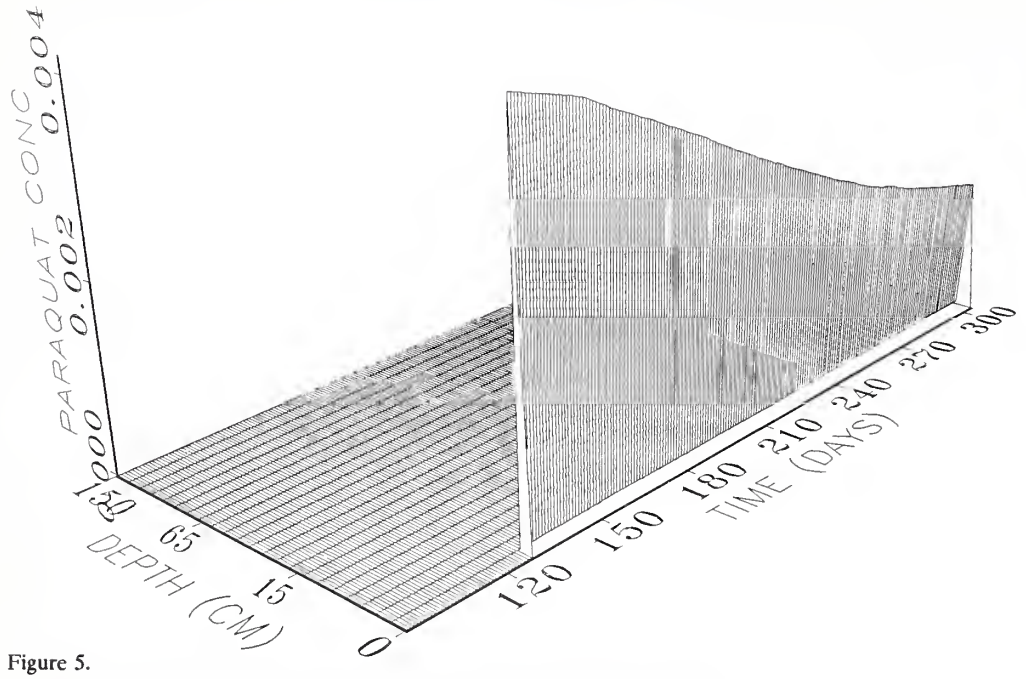


Figure 5.
Paraquat concentration in soil water after application of 1.1 kg/ha on day 126.

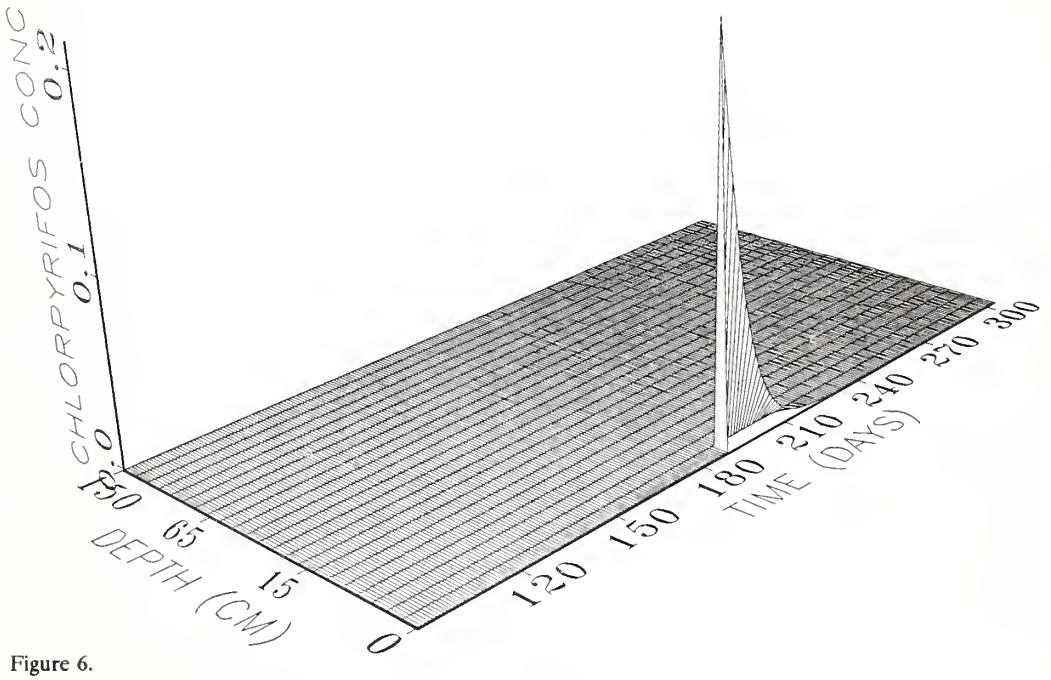


Figure 6.
Chlorpyrifos concentration in soil water after application of 4.4 kg/ha on day 190.

(6) the surface roughness created by its presence. This submodel consists of two major components, colonization (plant stage of growth) and production subsystems. The colonization or phenological growth stages can be used to show crop growth. Seven different stages are depicted. They are dormant seeds, germinating plants, emerging plans, juvenile plants, vegetative plants, flowering plants, and harvestable plants. Figure 7 shows the progression of corn plants, raised in the example of Figure 1, through the different phenological growth stages. The left axis is the

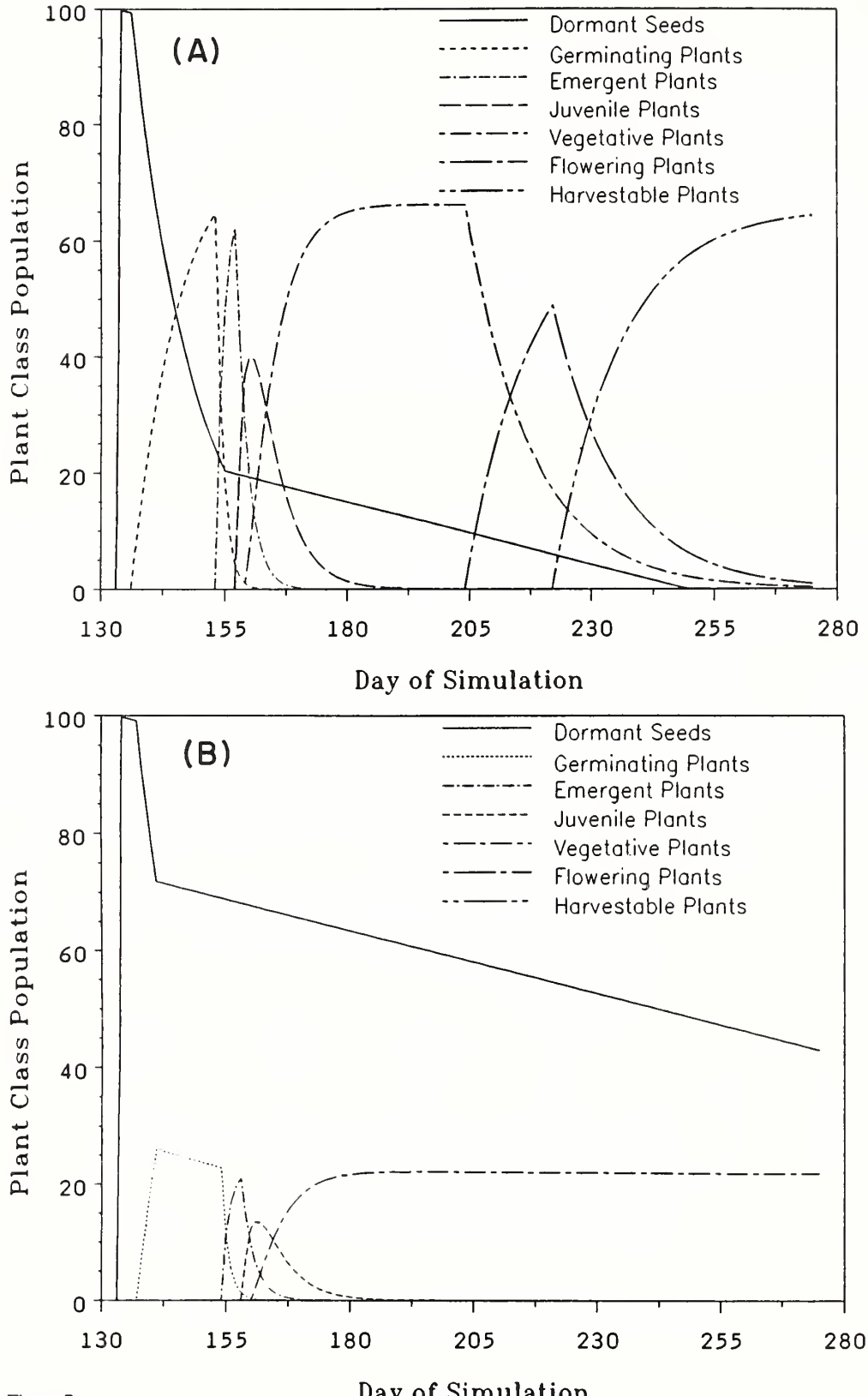


Figure 7.
Phenological growth stages of corn: A in an unstressed state of development and B stressed to the point of preventing flowering and development of harvestable material.

plant population by class as a percent. The plants shown in part A were essentially unstressed in the year we simulated. The populations in part B show he same curves under a stressed situation. Note the reduction in percent of plants in the latter stages of growth, the plants do not flower or show development of harvestable material.

Management Processes

The management submodel consists of a description of tillage and management processes defining the state of the root zone. It includes typical tillage practices for most common crop rotations and the impact these tillage practices have on surface roughness, erosivity, soil density, and micro and macroporosity. When not specified by the user, the timing of typical tillage practices (fertilizer and pesticide applications, irrigation and drainage, planting densities and timing, primary tillage, cultivation, and harvest operations) are assumed functions of soil moisture and crop conditions. Notill features are also included.

THE INPUT FILE GENERATOR

The input file generator assembles data in the format needed by the six major processes. It is designed to call information from pesticide, soils, management, plant, nutrient, and soil chemistry data bases; and relies, when necessary, on default values. It interrogates the user for site specific information and provides help in the form of questions and tables. Default values are provided automatically when necessary. Figure 8 is an example of one of the input situations showing overlays and a help screen. Flashing background and a blinking cursor show the user the data requested. The input file generator is designed to keep the model as invisible as the user desires it to be. In other words one entry on the user's part may enter all that is needed for some processes, pesticides for example. The user can observe, and change values of a given pesticide characteristic if he wishes, but he is cautioned not to do so, unless he has expertise in the area. The user can completely bypass many screens and doesn't even need to know they exist. It's designed to be as flexible as possible for both the experienced and inexperienced user.

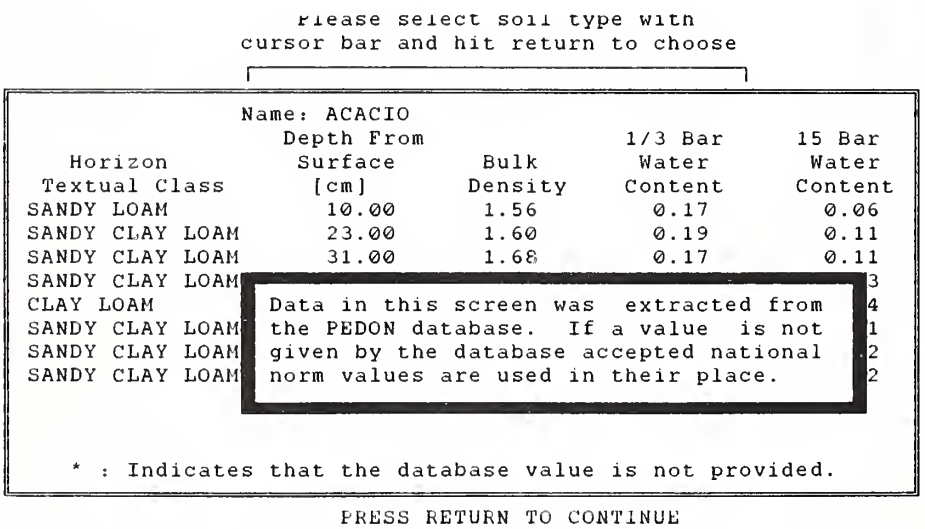


Figure 8.
A typical input screen a with help overlay

THE OUTPUT REPORT GENERATOR

The output report generator interrogates the user to determine the output desired, then assembles the information into a report with an easy to read format. It includes routines to summarize the data into user defined periods such as daily, monthly, and yearly totals. It also displays the information in graphical form if desired. Use of the output report writer is facilitated by a series of "canned" output packages. These include one for each of the major processes supplemented with hydrologic and other data. There is also one general output that includes selected features from each component. If the user desires something different, a tailor-made selection is possible. The figures shown in this paper are typical of the output received. Each of the "canned" output packages consists of about 8-10 different kinds of data. It is not possible to get much more than this because of the tremendous quantity of data that must be stored internally. Thus if several different scenarios are desired, it may be necessary to run the model several times. at present, the model running a full complement of fertilizer, 3 pesticides, crop growth, etc., for a growing season requires about 20 minutes on a 386 PC.

COMPUTER SYSTEM REQUIRED

Code for the simulation model conforms completely to ANSI-FORTRAN-77 programming standards and limitations. The Input Data Generator and Output Report Generator programs are designed specifically for use on an IBM-compatible microcomputer system. The model has been coded simultaneously for use on a 386 PC and a mini system such as a DEC MicroVax II.

REFERENCES

Nicks, A.D. and L.J. Lane. 1989. Weather Generator. Chapt. 2 in USDA Water Erosion prediction Project: Hillslope Profile Model Documentation. L.J. Lane and M.A. Nearing editors. USDA National Soil Erosion Laboratory, pp. 2.1-2.19.

DeCoursey, D.G., K.W. Rojas, and L.R. Ahuja. 1989. Potentials for non-point source groundwater contamination analyzed using RZWQM. ASAE Paper No. 892562, presented at 1989 International Meeting of ASAE. New Orleans, LA Dec. 12-15, 1989.

OPUS: AN ADVANCED SIMULATION MODEL FOR NONPOINT-SOURCE POLLUTION TRANSPORT AT THE FIELD SCALE -- AN OVERVIEW

V. A. Ferreira and R. E. Smith[1]

ABSTRACT

Opus is a computer simulation model of an agricultural system. The system watershed is a single field, describable as having laterally homogeneous soil horizons. Input rainfall is characterized by the record from a single gage. Processes modeled include surface and root zone hydrology, erosion, crop growth, evapotranspiration, nutrient cycling, pesticide processes, chemical transport, and agricultural management. The objective of Opus is to indicate system response relative to various management practices.

INTRODUCTION

The Opus computer program is the implementation of a comprehensive mathematical model of an agricultural field's response to environmental driving variables, under various management practices. After the publication of CREAMS (USDA, 1980), development of CREAMS2 began. It was designed to be a single program with considerably more feedback among components than CREAMS contained. The new model was made to be more responsive to management changes. Components were designed for better matching of technical complexity; for example, the daily hydrology option was paired with a similarly lumped erosion option. Restructuring, component replacements and modifications, team changes, and other major differences necessitated renaming the model, now called Opus. Opus is nearing release and is described and demonstrated here. The mathematical models incorporated in Opus are described in detail by Smith (in press).

OVERVIEW

The FORTRAN computer program named Opus is the implementation of mathematical models of physical and chemical processes on an agricultural field. The objective is to simulate relative hydrologic, erosion, and chemical fate results from various management and climate scenarios. Opus may be used in various types of agricultural studies, from simple management practice analysis to complex research applications. Two levels of hydrologic complexity are available; users must choose the level most compatible with their objectives and available input data.

Figure 1 shows the processes simulated by Opus. Arrows between processes indicate model feedback pathways. Required input includes numerical description of the field, including topography, soils, climate, initial conditions, and management practices. The user has numerous options regarding the complexity of input and output data. In many cases detailed input can be avoided by employing Opus default values, including functional relationships among variables. For example, a user interested in detailed soil moisture accounting may input parameters such as saturated hydraulic conductivity and bubbling pressure; or the user may simply input porosity and sand, silt, and clay fractions and Opus estimates reasonable values for the soil hydraulic parameters.

[1]Mathematician and Hydraulic Engineer, respectively. USDA-Agricultural Research Service, P.O. Box E, Ft. Collins, CO, USA 80522.

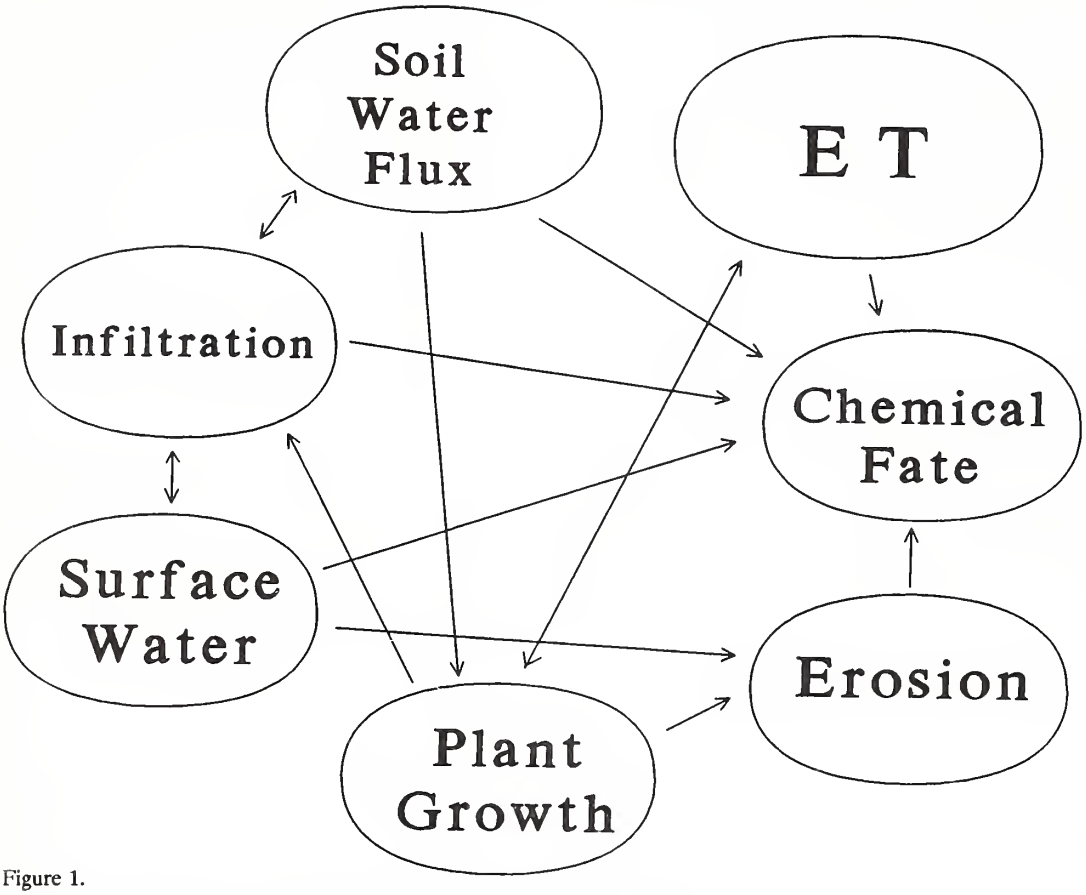


Figure 1.
Processes simulated by Opus, with feedback pathways indicated by arrows.

Opus operates on various time scales, depending on the process and on conditions within the simulation. As illustrated in figure 2, time scales vary from minutes to years. As with the complexity of input data, the simulation time scales are dependent upon user needs and data availability.

A major improvement over other models is the inclusion of Richardson's WGEN model (Richardson and Wright, 1984). This component generates daily values of solar radiation and maximum and minimum temperatures, correlated with rainfall occurrence; the user can, alternatively, input these values, if desired. For Curve Number simulations (daily hydrology), WGEN can generate reasonable sequences of daily rainfall data. Richardson derived input parameters for numerous U.S. locations; his maps and tables are included in the Opus User Manual (Ferreira and Smith, in press).

HYDROLOGY

The emphasis of Opus is on hydrology, because most pollutant movement is due to transport by water. Plant, weather, nutrient, and pesticide components are included to produce reasonable system simulation and to provide nonpoint-source pollution information.

Runoff

Two methods of runoff prediction are offered: an infiltration-based runoff model requiring breakpoint rainfall input, and an SCS Curve Number runoff model driven by daily rainfall data.

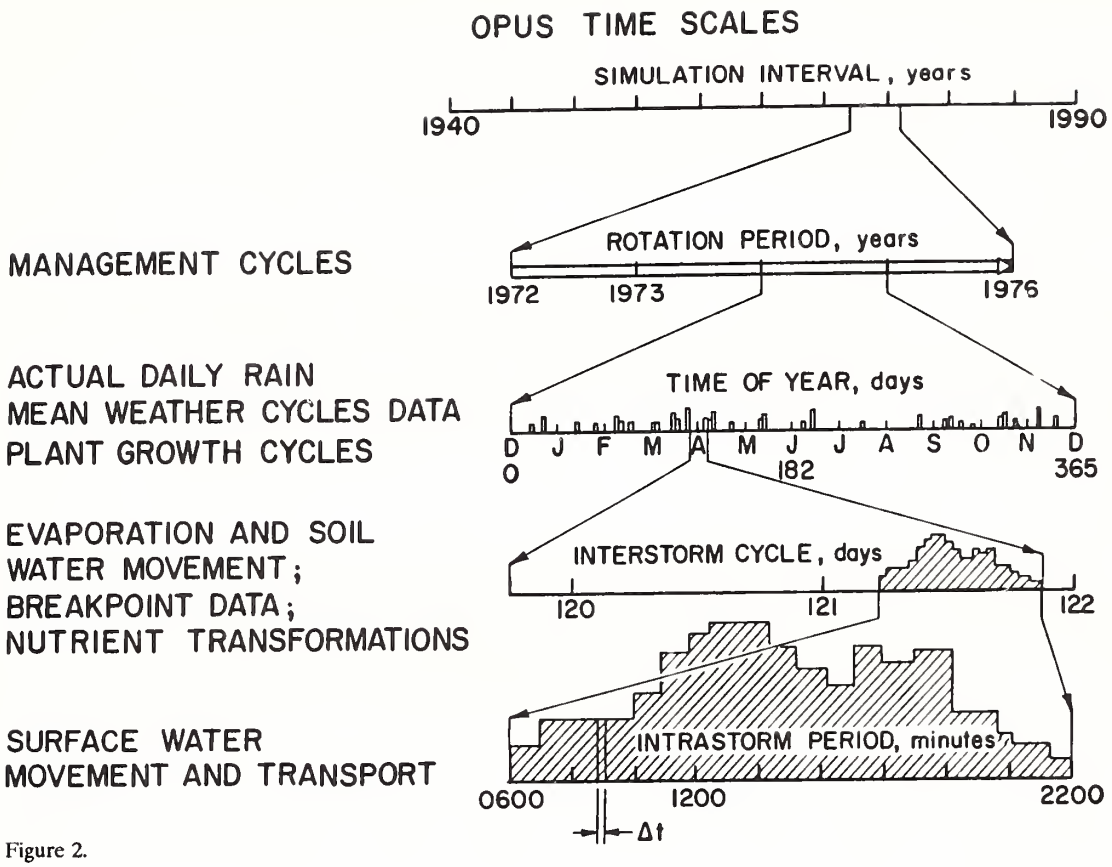


Figure 2.
Time scales of various Opus processes.

The Curve Number model yields daily runoff volume predictions. Runoff peaks are predicted from an empirical relationship which is a function of daily volume. Parameters of the peak/volume relation vary seasonally to account for seasonal variability of rainfall intensities. Daily rainfall information required to drive the Curve Number model may either be input or generated internally from input weather statistics. The infiltration model provides optional runoff hydrograph output.

Soil Water Movement

The scheme used for redistribution of soil water is one of the strongest points of the Opus model. Where many currently popular models use a "filled reservoir" system, Opus provides more physically realistic simulation with dynamic solution of Richards' equation. The solution technique used is adaptive, taking large time steps when the system approaches steady state and small steps when the system is in rapid transition. Small steps are needed after a storm or irrigation event, and particularly between layers with very different water-holding characteristics. This strategy combines physical realism with computational efficiency. The soil water model predicts interlayer fluxes in a manner that is well-suited to chemical transport simulations.

Figure 3 demonstrates an analysis of soil water information. It shows simulation results for part of a growing season in Watkinsville, GA. Soil water content is plotted versus time and depth, to show system (and model) performance. Days on which rainfall occurs appear as surface spikes which rapidly dissipate as the infiltrated water is redistributed in the soil. As plants develop throughout the growing season, the roots grow deeper and withdraw water due to transpiration.

Obvious breaks between layers indicate interfaces between horizons with different hydraulic properties.

Evapotranspiration

The evapotranspiration (ET) scheme of Opus is adapted from CREAMS. Driving variables are solar radiation and mean daily air temperature. First, a potential ET is calculated using a modified Penman approach; then actual evaporation and transpiration are estimated based on field conditions at the time.

EROSION

Opus has two erosion options which parallel its hydrology options. The simple model (daily) used with Curve Number hydrology is MUSLE, the USLE modification derived by Williams (1975). The more complex infiltration model is linked with a spatially distributed sediment transport component modified from KINEROS (Smith, 1981) to allow simulation of various particle size classes.

MANAGEMENT

Opus is designed to indicate the relative response of a field under various management strategies. Figure 4 illustrates the general management scheme in Opus. Options include land treatment, cropping, cultivation, planting, harvest, grazing, irrigation, chemical and manure applications, and plowing. Management is specified on a rotation basis to minimize necessary input. Most field operations are described by the user in terms of descriptive information (e.g., depth of plowing and chemical characteristics of pesticides) and target dates for management occurrence.

Land treatments simulated by Opus include terracing, inclusion of an impoundment, and use of grass buffer strips. Plowing to form furrows is assumed to control the direction of flow for specified periods.

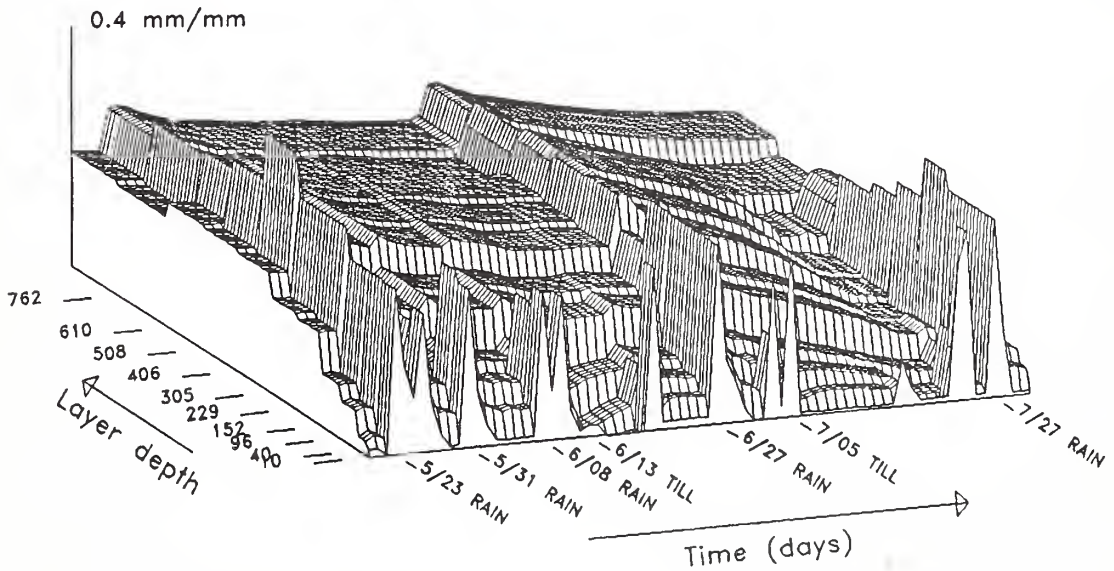


Figure 3.
Plot of soil moisture content versus time and depth (from Ferreira and Pons, 1989).

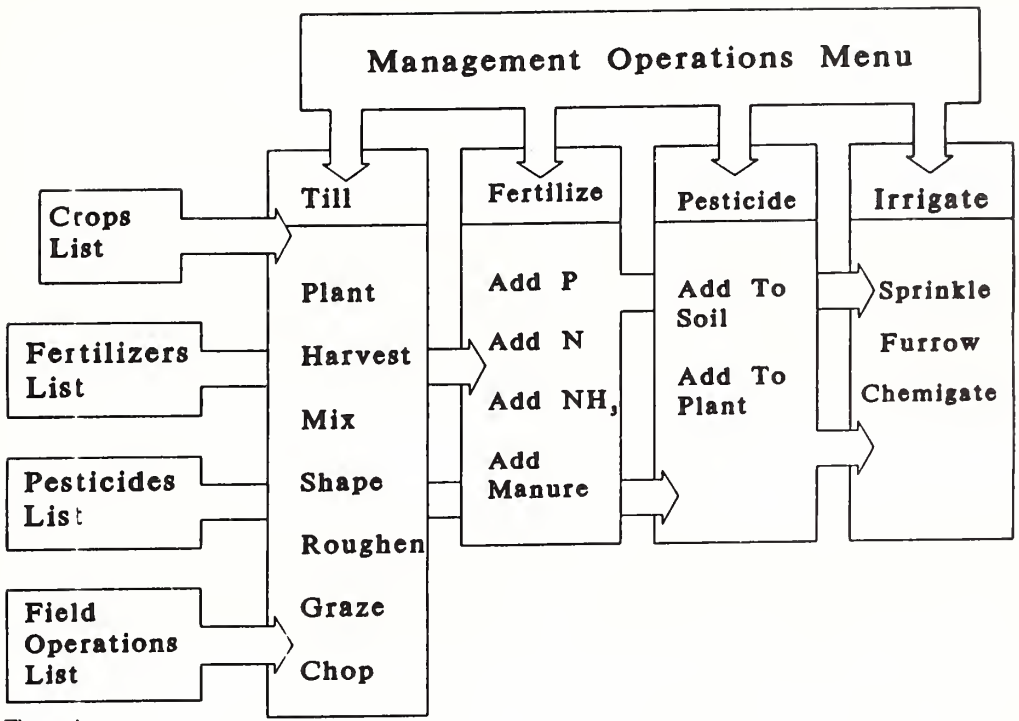


Figure 4.
Opus management scheme.

Crop management is a powerful component of Opus. As described later, the plant model is designed to simulate a wide variety of crops. Opus allows several crops each year, overlapping or in sequence, and is responsive to fallow periods as well.

Tillage and plowing operations are described in terms of surface roughness, mixing efficiency (used to redistribute chemicals), furrow height and width, and depth of the operation. Such mechanical operations affect not only the surface hydrology, but also other components including the nutrient cycle and pesticide fate. Seeds or seedlings may be planted to any specified depth. Row spacing is user-defined. Harvesting, including grazing, is described in terms of the amount of plant matter removed with the seed/fruit portion. Root crops can also be simulated by Opus.

Irrigation options include both furrow and sprinkler irrigation, either scheduled or "on demand" (whenever the soil moisture status reaches a threshold). Irrigation advance and irrigation efficiency are simulated by surface water hydraulics.

Chemical applications (fertilizer and pesticides) may be aerial, plant- or soil-directed, by injection into the soil, or in irrigation water. Atmospheric losses of pesticides are considered. Nutrients may be applied as either chemical fertilizer or manure.

CROP GROWTH

The plant component is mechanistic, driven by air temperature and solar radiation. Plants are stressed by water or nutrient deficiencies and temperature. Daily plant mass accumulations are allocated to roots, fruit, and above-ground biomass (leaves and stems). Both annual and perennial plants may be simulated. Figure 5 demonstrates the response of a perennial grass to the Georgia climate under irrigated conditions. Root, leaf, and seed fractions are shown, as are two harvests

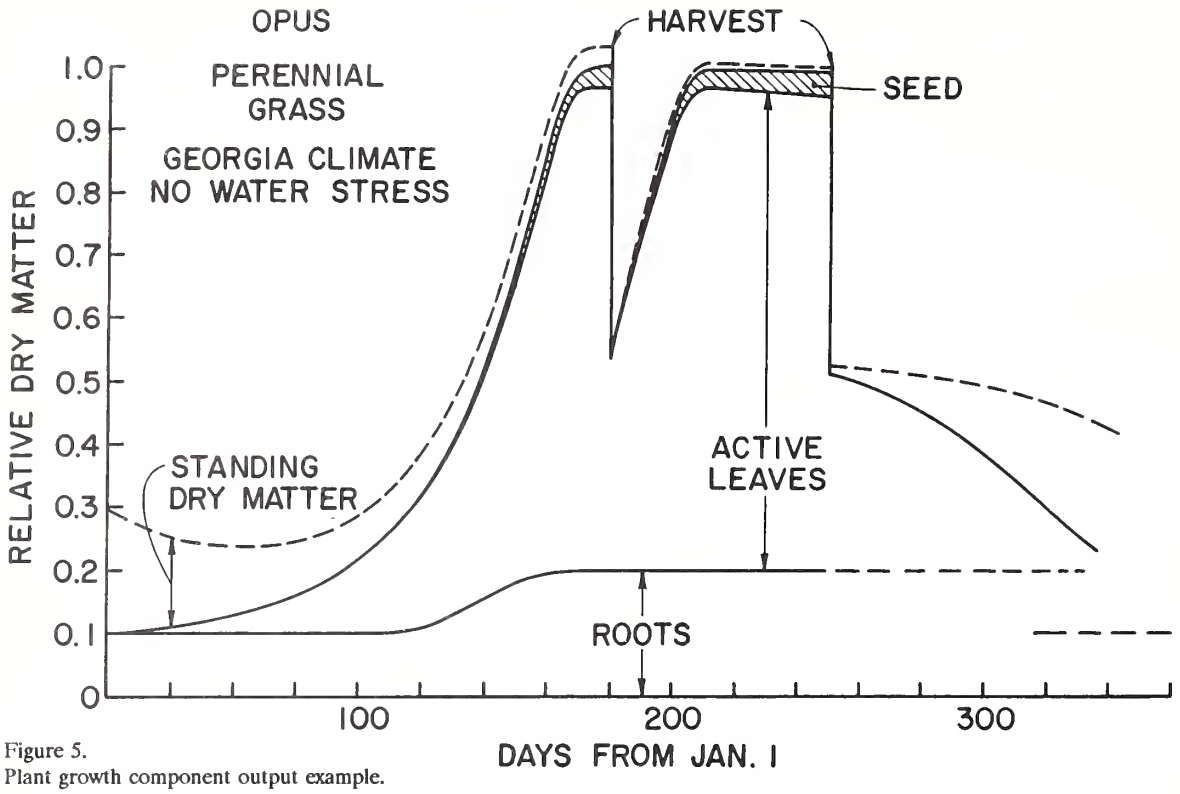


Figure 5.
Plant growth component output example.

and a period of senescence. Plants enter such a period after maturity, during which leaf material is slowly shed. The model has performed well for a wide variety of agronomic crops, including corn, wheat, soybeans, and range grasses.

AGRICULTURAL CHEMICALS

One objective in the development and maintenance of Opus is to provide insight into the relationship between agricultural management and nonpont-source pollution. The chemical components are therefore comprehensive, including a considerable amount of interaction with the environment and sensitivity to management practices. The two categories of chemical components are pesticides and nutrients. Application, dissipation, and transport are the main processes governing chemical fate. Many subprocesses are also simulated by Opus.

Nutrient Cycle

Parton adapted the Century model (Parton et al., 1987) for Opus. Nitrate-N and phosphate-P are accounted for, in a carbon-based system. Processes simulated include nitrification, denitrification, mineralization, immobilization, and plant nitrogen fixation. The model has been tested most extensively on grasslands, but has performed well in other agricultural situations, including cases of heavy manure applications. It has operated with noteworthy stability for long-term Opus simulations (on the order of 50-100 yr).

Pesticides

The pesticide component is an improved version of that in CREAMS. One improvement is in relating the adsorption coefficient, Kd, to the organic carbon content of each soil layer. This allows better simulation in agricultural settings where organic matter is often concentrated in

surface layers, trapping adsorbed pesticides. A kinetic option has been added to account for the slow desorption of some chemicals, for which the traditional equilibrium models perform poorly. Figure 6 demonstrates the relative distributions of a mobile pesticide for days 8, 17, and 174 after application for the equilibrium and kinetic cases. The equilibrium assumption causes the pesticide to move down from the surface much more quickly than in the kinetic case. The gross difference of surface concentration on day 174 indicates that resulting concentrations in runoff and with eroded sediments will be profoundly different for the two cases, although all other chemical attributes were identical.

INPUT/OUTPUT

Opus is designed to optimize the trade-off between complexity and simplifications in input, simulation, and output. In many cases, only minimal input data are required, and the user has the option to enter more detailed information, when available. When applicable, default values are provided. The User's Manual discusses the physical meaning of each input variable and provides guidance as to ranges of reasonable values under various conditions.

Input

Opus input is designed to be much like CREAMS input, because of the number of users familiar with it. Several improvements have been made to facilitate creating, modifying, and finding errors in input files. Figure 7 is an excerpt of a parameter file. The file is in a template format rather than a solid file of numbers. A variable name is above each value, so the user will not change, for example, the watershed area when meaning to change the Curve Number. This type of error has been common in usage of programs without templates. The identifying code in the far right-hand column (e.g., B01) is used by the program to check that the correct record is being read. This solves a common problem in input files with a variable number of records, where the program can mistake, for example, pesticide input for initial conditions. Opus performs numerous error checks

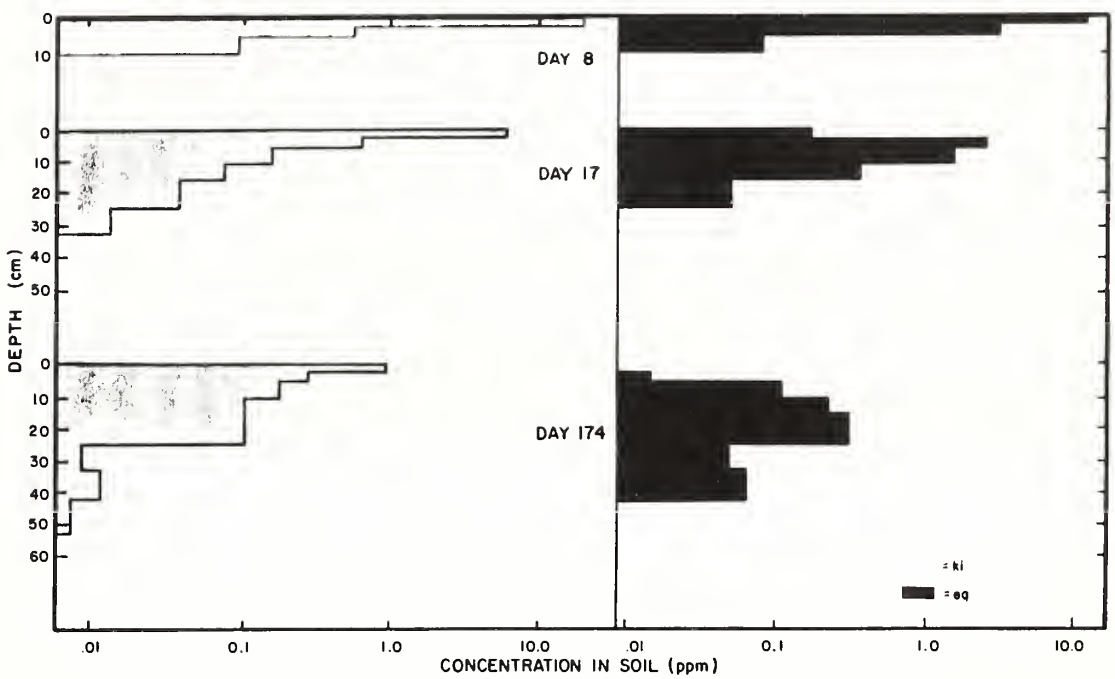


Figure 6.
Pesticide concentration distribution with soil depth for kinetic (ki) and equilibrium (eq) adsorption models. From Smith and Ferreira (in press).

```
**************************************************************************
**                                                                      **
**  G E N E R A L    W A T E R S H E D  &  I N I T I A L    S T A T E  **
**                                                                      **
**************************************************************************
    DA    DWTB    GLAT     CN2    PHRN   CONRN    DASL    ALBS   FWIND
 3.11    10.0    31.0     80.    6.00    0.80    0.39    0.35    0.28            B01
SRESD   STDRY    THST   ROWSP    DPFR RGSURF     ZSF   DTILL
100.0   100.0    0.20     0.5     3.0     0.5     0.5      0.                    B02
-------------------------------------
 There must be NC values of ICR on B03
-------------------------------------
   NC     ICR     ICR     ICR     ICR
    1       1                                                                  B03
**********************************************************************
**                                                                  **
**                S O I L    H O R I Z O N    D A T A               **
**                                                                  **
**********************************************************************
   NSL
    2                                                                         C01
 ----------------------------------------
  There must be NSL sets of C02-C03 prompts
  and records, one for each horizon
  ----------------------------------------
  GZH     POR   PSAND   PSILT   PCLAY      RC     B15    PBUB    ALAM     THS
 6.00     .41     .64     .19     .17    0.55     .13     4.9       0     .33       C02
 ORGC   SRSDU    WNO3   WPLAB     SPH     PKD   FEROD     OMN     OMP   TOTP
  .61    100.   250.0    22.0     5.9    -1.0    0.00    -1.0    -1.0   -1.0       C03
  GZH     POR   PSAND   PSILT   PCLAY      RC     B15    PBUB    ALAM     THS
 60.0     .49     .51     .27     .22    -1.0     .28       0       0     .45       C02
 ORGC   SRSDU    WNO3   WPLAB     SPH     PKD   FEROD     OMN     OMP   TOTP
  .05      2.     20.      1.0     6.0    -1.0     0.0    -1.0    -1.0   -1.0       C03
 **********************************************************
 **                                                      **
 **           C R O P        D A T A                    **
 **                                                      **
 **********************************************************
 NCROP
    1                                                                         D01
 -------------------------------------
  There must be NCROP sets of D02-D04
  prompts and data records
 -------------------------------------
  IDCR    IPER    PLAI    DDEM    DDMX   PDRYM    POTY     RDP    PLIG    RLIG
soybean      0    4.00   150.0   3600.   9000.   2300.    24.0    0.15    0.10       D02
 POTHT    PPCV    TGBM    TGOP   CONVF   DEACT    COVI DMINIT
   4.0    0.60    40.0    90.0    20.0    0.02     0.0     0.0                     D03
  CONY    CFXN     PNO     PNF     DKC   PNRAT
 0.065     1.0    0.04   0.018    3.50    0.25                                   D04
```

Figure 7.
Excerpt from input parameter file.

as the data are read, including checking for reasonable values of individual parameters as well as in relationships among parameters.

Output

Opus generates user-specified output as needed for various analyses, not overwhelming either the user or the computer with unnecessary output. Several output options are available, representing various time and space scales. The user chooses output scales and frequency based on the process of interest. For example, runoff predicted by the Curve Number method is output in monthly summary tables, illustrated in figure 8. The user may request optional output daily on storm days, as shown in figure 9. This option includes information about simulation start and end dates, and indicates dates of management operations. Infiltration model predictions may be output as monthly or storm information, but also may be viewed within the storm, as shown in figure 10. This option is controlled by a user-input threshold rainfall value, below which output is not written. This allows the user to view only the larger storms, if desired, and can avoid large masses

ANNUAL SUMMARY FOR CROP YEAR 1973

	PRECIP MM	RUNOFF MM	TOT E.T. MM	SEEP MM	SOIL WATER MM	SEDMT. T/HA	NO3 STRESS	DAYS	H2O STRESS	DAYS
					230.					
MAY	166.	25.	72.	66.	234.	0.09	0.0	0.0	0.	0.0
JUN	121.	15.	91.	32.	217.	0.47	0.0	0.0	0.	0.0
JUL	123.	28.	101.	3.	207.	0.66	0.0	0.0	0.	0.0
AUG	23.	0.	47.	-9.	193.	0.00	0.0	0.0	22.	25.1
SEP	135.	19.	92.	2.	216.	0.30	0.0	0.0	7.	8.1
OCT	5.	0.	32.	-8.	196.	0.00	1.0	1.0	20.	19.9
NOV	43.	0.	20.	-6.	226.	0.00	0.0	0.0	15.	15.6
DEC	183.	11.	23.	84.	281.	0.44	0.0	0.0	0.	0.0
TOT	799.	98.	479.	162.		1.97				

Figure 8.
Excerpt from Opus standard summary output file.

DATE	RAIN	RUNFF	PEAK DISCH	PERC	AVG TEMP	SOIL MOIS	TRANS	EVAP	SED YIELD	ENR RAT	MANAGEMENT OPERATIONS
MMDDYY	MM	MM	MM/HR	MM	DEGC	MM/MM	MM	MM	T/HA		
043073											SIMULATION START
050273	10.			0.	18	.310	0.	4.			
050873	16.			2.	18	.314	0.	11.			
051973	6.			3.	22	.299	0.	14.			
051973	7.			0.	22	.307	0.	0.			
051973	14.	0.0	0.0	0.	22	.325	0.	0.	0.00	0.23	
052273											FERTILIZED
052273											DISK hr
052373	22.			7.	19	.329	0.	12.			
052873	48.	14.7	4.6	9.	23	.345	0.	13.	0.06	0.16	
052873	43.	10.3	3.4	8.	23	.374	0.	3.	0.03	0.16	
060173	6.			37.	27	.314	0.	15.			
060473											FERTILIZED
060473											rolcult
060573	7.			1.	27	.302	0.	15.			
060573	5.			0.	27	.306	0.	2.			
060673	39.	14.6	45.4	0.	26	.337	0.	2.	0.43	0.20	
060773	20.	0.4	4.0	3.	27	.355	0.	3.	0.02	0.20	

Figure 9.
Excerpt from option stormwise summary output.

of unnecessary output. The final output option available is subsurface information, as illustrated in figure 11. The user controls the frequency of this output (monthly, yearly, or every n days, n being user-specified).

MODEL STATUS

The two-volume Opus documentation is scheduled to be published by the USDA-ARS in 1990. The computer program will then also be available. Opus has been tested under many conditions, including various crops, climates, and soil types, under an assortment of management strategies. Some of the testing and validation is reported in the model documentation; other results have already been reported in the literature (e.g., Smith et al., 1986; Ferreira and Pons, 1989).

REFERENCES

Ferreira, V.A. and F.A. Pons. 1989. Hydrologic Subprocess Model Effect on Chemical Fate Predictions. Amer. Soc. Ag. Engrs. Winter Meeting, Paper No. 89-2523. 22pp.

HYDROGRAPH FOR RAINFALL OF 25.40 MM ON APR 13 74:
FURROWED SURFACE FLOW CONFIGURATION

TIME MIN	RAIN RATE MM/HR	INFIL RATE MM/HR	PLANE DISCHRG MM/HR	OUTLET DISCHRG MM/HR	ACCUM RUNOFF MM	SEDIMENT DISCHRG KG/MIN
184.0						
191.0	10.	10.	0.	0.	0.	0.00
195.4	64.	61.	0.	0.	0.	0.00
196.4	64.	55.	0.	0.	0.	0.00
197.	64.	51.	1.	0.	0.	0.01
197.5	51.	42.	1.	0.	0.	0.05
198.2	51.	40.	2.	0.	0.	0.21
198.9	51.	38.	3.	0.	0.	0.48
230.5	4.	21.	1.	11.	1.9	25.68
231.1	4.	21.	0.	11.	2.0	25.45
232.2	4.	21.	0.	10.	2.2	24.67
235.5	4.	4.	0.	7.	2.6	22.82
240.5	4.	4.	0.	3.	2.8	13.58
250.7	4.	4.	0.	0.	2.9	0.06

SUMMARY OF SURFACE RUNOFF FOR APR 13 74:
RUNOFF OF 4.8 MM, SEDIMENT YIELD 0.435 T/HA

NUTRIENTS IN RUNOFF: GM/HA NO3= 65.58
NH4= 753.6
PO4= 78.51

PESTICIDES AND RESPECTIVE RUNOFF AMOUNTS IN GM/HA:

TRIFLUR,	2.e-06	WITH SEDIMENT,	2.5e-03	DISSOLVED
PARAQ,	3.9	WITH SEDIMENT,	3.4e-02	DISSOLVED
DIPHENA,	0.0	WITH SEDIMENT,	3.e-06	DISSOLVED

Figure 10.
Sample within-storm surface hydrology output.

Ferreira, V.A. and R.E. Smith. In press. Opus: An Advanced Simulation Model for Nonpoint-Source Pollution Transport at the Field Scale. Vol. II. User Manual. To be published by the USDA Agricultural Research Service.

Parton, W.J., D.S Schimel, C.V. Cole, and D.S. Ojima. 1987. Analysis of factors controlling soil organic matter levels in Great Plains grasslands. Soil Sci. Soc. Am. J. 51(5):1173-1179.

Richardson, C.W. and D.A. Wright. 1984. WGEN: A model for generating daily weather variables. United States Dept. of Ag., Agricultural Research Service, ARS-8, 83 pp.

Smith, R.E. 1981. A kinematic model for surface mine sediment yield. Trans. Am. Soc. Ag. Eng. 24(6):1508-1514.

SUBSOIL STATUS AS OF APR 30 73 :

LAYER NO.	DEPTH (MM) FROM:	TO:	TEMP. DEGC	WATER CONTENT	CAPILLARY HEAD, MM
1	0.0	10.0	17.1	0.200	-1668.
2	10.0	40.0	17.1	0.200	-1668.
3	40.0	96.2	17.1	0.200	-1668.
4	96.2	152.4	17.1	0.200	-1668.
5	152.4	228.6	17.1	0.292	-1668.
6	228.6	304.8	17.1	0.292	-1668.
7	304.8	406.4	17.1	0.339	-1668.
8	406.4	508.0	17.0	0.339	-1668.
9	508.0	609.6	17.0	0.339	-1668.
10	609.6	762.0	17.0	0.345	-1668.
11	762.0	952.5	16.9	0.345	-1668.

MOBILE NUTRIENTS IN ROOTZONE LAYERS:

LAYER NO.	DEPTH (MM) FROM:	TO:	NITRATE [KG/HA]	AMMONIA [KG/HA]	PHOSPHATE [KG/HA]
1	0.0	10.0	11.7	27.4	3.4
2	10.0	40.0	35.2	82.1	10.3
3	40.0	96.2	65.9	153.8	19.3
4	96.2	152.4	65.9	153.8	19.3
5	152.4	228.6	66.6	155.5	13.3
6	228.6	304.8	66.6	155.5	13.3
7	304.8	406.4	20.6	48.1	1.4
8	406.4	508.0	20.6	48.1	1.4
9	508.0	609.6	20.6	48.1	1.4
10	609.6	762.0	12.4	28.8	10.7
11	762.0	952.5	15.4	36.0	1.5

PESTICIDE TOTALS IN ROOT ZONE LAYERS:

LAYER NO.	DEPTH (MM) FROM:	TO:	TRIFLUR GM/HA	PARAQ GM/HA	DIPHENA GM/HA
1	0.0	10.0	135.	6170.	1120.
2	10.0	40.0	4.04	185.	33.6
3	40.0	96.2	0.756e-01	3.46	0.630
4	96.2	152.4	0.756e-03	0.346e-01	0.630e-02
5	152.4	228.6	0.103e-04	0.470e-03	0.854e-04
6	228.6	304.8	0.103e-06	0.470e-05	0.854e-06
7	304.8	406.4	0.137e-08	0.626e-07	0.114e-07
8	406.4	508.0	0.00e+00	0.00e+00	0.00e+00
9	508.0	609.6	0.00e+00	0.00e+00	0.00e+00
10	609.6	762.0	0.00e+00	0.00e+00	0.00e+00
11	762.0	952.5	0.00e+00	0.00e+00	0.00+00

Figure 11.
Sample subsurface output.

Smith, R.E. In press. Opus: An Advanced Simulation Model for Nonpoint-Source Pollution Transport at the Field Scale. Vol. I. Model Documentation. To be published by the USDA Agricultural Research Service.

Smith, R.E. and V.A. Ferreira. In press. Effect of soil adsorption kinetics on agricultural pesticide fate, submitted J. Environ. Sci. and Health, Part B. Pesticides, Food Contaminants, and Agricultural Wastes.

Smith, R.E., R.M. Summer, and V.A. Ferreira. 1986. Model feasibility study of atmosphere to surface water radioactive pathways. Prepared for U.S. Nuclear Regulatory Commission, Interagency Agreement No. NRR-85-066, FIN No. D1335, B&R No. 20-19-50-53-2. 27pp.

USDA. 1980. CREAMS: A field scale model for chemicals, runoff, and erosion from agricultural management systems. W.G. Knisel, ed. USDA Conservation Research Report No. 26, 643 pp.

Williams, J.R. 1975. Sediment yield prediction with universal equation using runoff energy factor. United States Dept. of Ag., Ag. Research Service ARS-S-40: 244-252.

A SIMPLE METHOD FOR ASSESSING PESTICIDE LEACHABILITY

David I. Gustafson[1]

ABSTRACT

A simple index for predicting pesticide leachability is developed based on graphical examination of a plot formed by two widely available pesticide properties: half-life in soil (t1/2) and partition coefficient between soil organic carbon and water (KOC). Other physical properties have been advanced as indicators of leachability, but they are shown to have no useful power in discriminating between "leachers" and "non-leachers." Scores assigned with the new screening index agree with the results of several recent well-water monitoring programs. A nomogram is given that reduces the task of calculating the index to simply placing a straightedge on a diagram.

INTRODUCTION

Several authors have proposed screening methods for determining whether a pesticide is likely to leach to groundwater in detectable quantities. Some methods use threshold values for a physical property or set of properties which, when exceeded, indicate that the chemical will leach (Wilkerson and Kim 1986, Cohen et al. 1984). Others have proposed simplified analytical or numerical solutions to the convective-dispersive equation using the measured or estimated properties of the chemical in order to predict the likelihood of leaching (Rao et al. 1985, Jury 1987, Enfield 1982, Carsel 1984, Dean 1984).

A different approach is taken in this paper. Groups of compounds are examined that have been categorized with respect to their leachability, and for which consistent sets of physical properties have been collected. Properties of the chemicals representing soil mobility and soil persistence are plotted by leachability class to define a region that contains the leaching compounds.

As used in this paper, the term leachability refers to the following pesticide property: that when used in a normal agricultural manner under conditions conducive to movement, it moves down through the soil in quantities sufficient to be detected in nearby wells of proper construction. Areas with direct connection between the surface and saturated zones, e.g. agricultural drainage wells or sink holes, are excluded in this definition.

MOBILITY IN SOIL

Assuming linearity, the ratio of concentrations in the soil (Cs) and aqueous (Cw) phases is denoted as KD (Lyman 1982):

$$K_D = Cs/Cw. \qquad [1]$$

Obviously, those chemicals with higher KD values move more slowly through soil, because a higher fraction of the chemical is in the immobile soil phase at any time.

The KD values measured for a particular chemical on a range of different soils often vary proportionately with the organic carbon content (fOC) of the soil (Karickhoff 1984). This

[1]David I. Gustafson, Engineering Specialist, Monsanto Agricultural Company, St. Louis, MO.

relationship has been used to derive a measure of mobility in which the effects of soil type and management history are specifically accounted for:

$$K_{OC} = K_D/f_{OC}. \quad [2]$$

The advantage of K_{OC} is that it represents a soil-independent measure of compound mobility causing the resulting K_D to vary with soil conditions. It is therefore useful in making comparisons of mobility among pesticides.

MEASURING PERSISTENCE IN SOIL

Although pesticide degradation is rarely a strictly first-order process, such kinetics are nearly always assumed. In the case of leachability, the half-life of interest is the time it takes, in the field, for soil residues of the parent molecule to decline by 50%, denoted in this document as t1/2.

A METHOD OF COMBINING PERSISTENCE AND MOBILITY MEASURES

Rather than depending on mathematics and computer models to define the way physical property characteristics interact to determine leachability, the physical property data themselves were used. The selected set of data is the list compiled by the California Department of Food and Agriculture (CDFA) in producing its recent document on Specific Numerical Values (Wilkerson and Kim 1986). These data were used because the methodology developed herein hinges on two factors: a consistent set of physical properties for a group of pesticides and a consistent classification of the same pesticides as "leachers" or "non-leachers" following normal agricultural use (exclusive of point-source events). These criteria were used by the CDFA to develop its list.

Shown in figure 1 are the 22 "leachers" and "non-leachers" for which the CDFA was able to find average values for both K_{OC} and t1/2. Table 1 serves as the key for this figure -- it lists all 44 compounds classified by the CDFA and the K_{OC}'s and t1/2's that were assigned to each. The logarithmic scales are in base 10. As expected, the leachers occupy the left and upper portions of the figure, corresponding to pesticides which are more mobile and more persistent in soil. The graph itself is a perfectly reasonable way to distinguish the leachers from the non-leachers, but, given the desirability of a numerical index, the curved nature of the region containing the leachers suggests that a hyperbolic function should discriminate between the two classes of compounds. The two solid curves in this figure are hyperbolas generated by two values of the following function:

$$GUS = \log10(t1/2)[4 - \log10(K_{OC})]. \quad [3]$$

The two values of GUS (Groundwater Ubiquity Score) in the figure are 2.8 and 1.8: these appear to bracket the region in which transition occurs from leachers to non-leachers. A sensitivity analysis was used to show that such a one-unit transition region is sufficient to maintain proper classification in spite of uncertainty in the average values for t1/2 and K_{OC} (Gustafson 1989).

In practice, the three zones of figure 1 could be used in the following way. Compounds which fall in the leacher or transition zone would require further investigation by more sophisticated modeling technologies. The level of modeling effort required to show that a leacher-zone pesticide (GUS>2.8) is really not a "problem-compound" would be somewhat greater than that for a chemical in the transition zone (1.8>GUS>2.8). Compounds which fall in the non-leacher zone could safely be exempted from further consideration as possible "leachers." However, it should be remembered that point-source contamination of wells is always a possibility when chemicals are used carelessly or wells are improperly installed in areas of pesticide use and/or storage.

Table 1.
Physical Properties and Classifications Assigned by the CDFA: KOC is given in (ml/g) and t1/2 is given in (days) (N.D. indicates No Data or data did not meet CDFA Standards).

Leachers				Non-Leachers				Transition			
Key	Pesticide	KOC	t1/2	Key	Pesticide	KOC	t1/2	Key	Pesticide	KOC	t1/2
1	Aldicarb	17	7	17	Aldrin	N.D.	10	39	Alachlor	161	14
2	Atrazine	107	74	18	Chloramben	N.D.	N.D.	40	Carbaryl	423	19
3	Cyanazine	N.D.	14	19	Chlordane	19269	37	41	Carbofuran	55	37
4	DBCP	40	N.D.	20	Chlorothalonil	1380	68	42	Dieldrin	12100	934
5	DCPA	N.D.	100	21	Chlorpyrifos	6085	54	43	Dinoseb	5900	30
6	Dichloropropene	955	N.D.	22	2,4-D	53	7	44	Ethoprop	26	63
7	Diuron	389	188	23	1,3-D	68	N.D.	45	Fonofos	5105	25
8	EDB	78	N.D.	24	DDD	45800	N.D.				
9	Metolachlor	99	44	25	DDT	213600	38200				
10	Metribuzin	N.D.	37	26	Dicamba	511	25				
11	Naled	133	N.D.	27	Endosulfan	2040	120				
12	Oxamyl	26	8	28	Endosulfan Sulfate	N.D.	N.D.				
13	Picloram	26	206								
14	Prometon	577	N.D.	29	Endrin	11188	2240				
15	Prometryn	614	94	30	Heptachlor	13330	109				
16	Simazine	138	56	31	Lindane	1727	569				
				32	Pendimethalin	N.D.	N.D.				
				33	Phorate	1660	38				
				34	Propachlor	794	4				
				35	Silvex	N.D.	22				
				36	Toxaphene	7950	83				

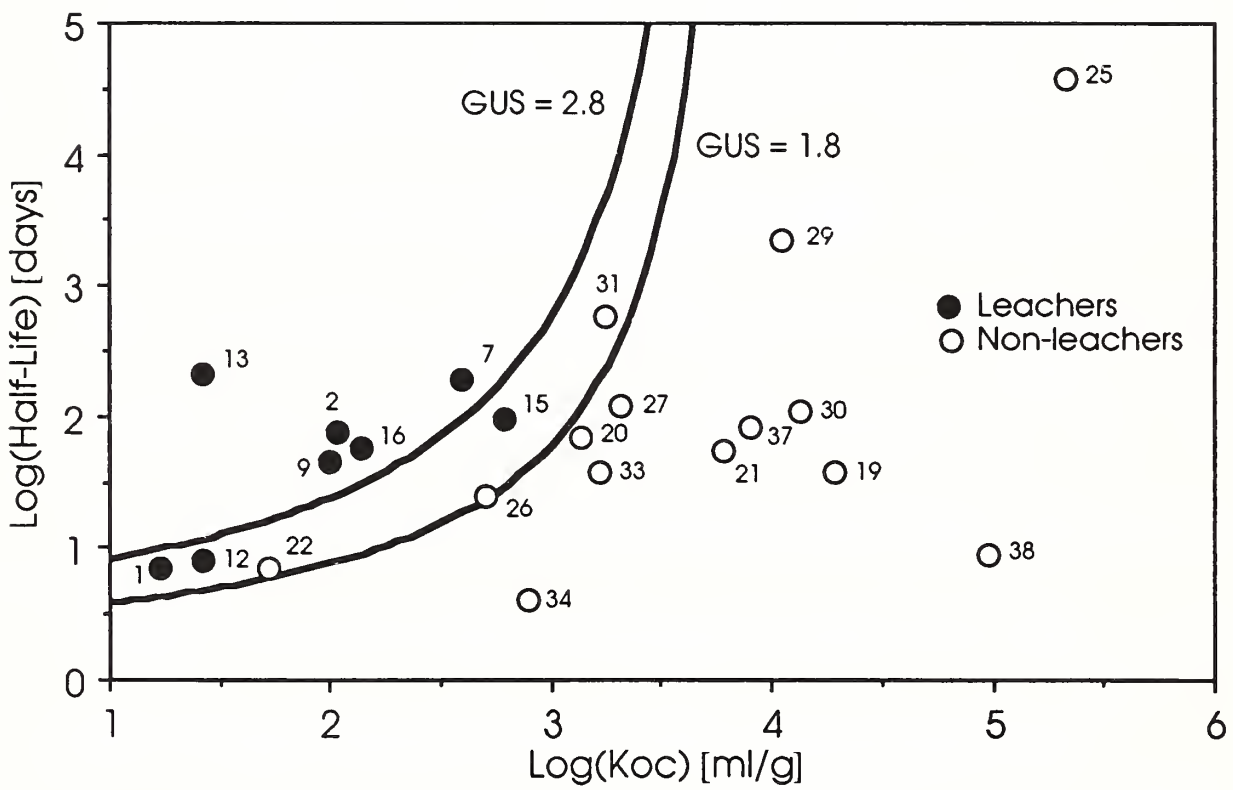


Figure 1.
Pesticides by key number (table 1) classified by the CDFA.

UTILITY OF OTHER PHYSICAL PROPERTIES AS LEACHING INDICATORS

An attempt was made to see whether any other physical properties might enhance the classification power of the GUS function (Gustafson 1989). There are fairly strong correlations between either water solubility or octanol-water partition coefficient and the GUS function, but there does not appear to be any useful additional separation of the leachers and non-leachers. Comparisons with estimates of pesticide volatility, based on a regression equation developed by Ralph Nash of the USDA,(Nash 1988) did not provide any additional classification power either.

COMPARISON WITH MONITORING DATA

A stringent test of the screening method is to compare its classifications with the results of well-water monitoring programs in which both nonpoint- and point-source pollution may be occurring. This was examined in some detail (Gustafson 1989) and excellent agreement was seen, despite the alleged predominance of point-source events in many incidents of contamination. This is probably because when spills or other mis-handling incidents occur near wells, the immobile materials are still prevented from showing up in the water due to strong binding and subsequent dissipation within the upper layers of the contaminated soil. An exception to this observation would be point-source incidents in which the pesticide is placed directly into the well (such as with back-siphoning).

A NOMOGRAM FOR CALCULATING GUS VALUES

The GUS value is easy to calculate from the soil persistence and soil mobility of the pesticide, assuming one has ready access to a calculator or computer with a base 10 logarithm function. If such computing "power" is not available, then the nomogram in figure 2 may be used. A

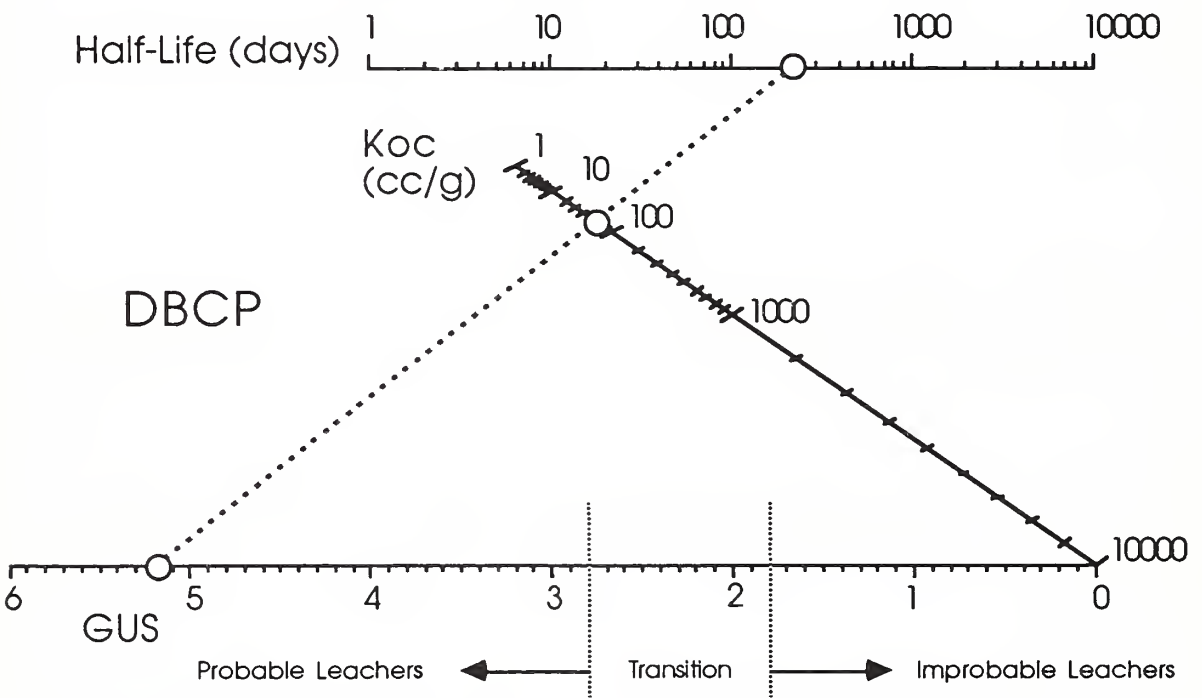


Figure 2.
Example use of the nomogram for calculation of GUS.

straightedge is lined up along the half-life and soil-water partition coefficient values for the chemical, and the GUS value is then simply read off the bottom scale. The example in figure 2 is for DBCP, a nematocide for which use was discontinued after it was found in thousands of wells in California (Wilkerson and Kim 1986, Jury 1987). Average physical properties for DBCP were calculated from all available literature data (Monsanto 1988).

A dynamic version of this nomogram has been developed on a Macintosh (II or SE) personal computer, upon which the straightedge in figure 2 may be dragged with a "mouse" by one of its handles along either the t1/2 or the Koc scale. This dynamic tool is particularly useful for investigating the effects of uncertainty in one or both of the physical properties making up the index. Readers interested in obtaining the program or the Pascal source code may contact the author.

ACKNOWLEDGMENTS

Larry R. Holden, a Fellow with Monsanto Agricultural Company, provided numerous suggestions on the statistical methodologies employed in this work, and his assistance is greatly appreciated.

LITERATURE CITED

Carsel, R.F., C.N. Smith, L.A. Mulkey, and others. 1984. Users Manual for the Pesticide Root Zone Model (PRZM): Release I. U.S. Environmental Protection Agency, Athens, GA. U.S. Government Printing Office: Washington, D.C., EPA-600/3-84-109.

Cohen, S.Z., S.M. Creeger, R.F. Carsel, and C.G. Enfield. 1984. Potential for pesticide contamination of groundwater resulting from agricultural uses. In R.F. Kruger and J.N. Seiber, eds., Treatment and Disposal of Pesticide Wastes. ACS Symposium Series no. 259. American Chemical Society, Washington, DC, pp. 297-325.

Dean, J.D., P.P. Jowise, and A.S. Donigian, Jr. 1984. Leaching Evaluation of Agricultural Chemicals (LEACH) Handbook U.S. Environmental Protection Agency, Athens, GA. U.S. Government Printing Office: Washington, D.C.; EPA-600/3-84-068.

Enfield, C.G., R.F. Carsel, S.Z. Cohen, and others. 1982. Approximating pollutant transport to ground water with PESTAN. Ground Water. 20:711-722.

Gustafson, D.I. 1989. Groundwater Ubiquity Score: A simple method for assessing pesticide leachability. Journal of Environmental Toxicology and Chemistry. 8:339-357.

Jury, W. A., D.D. Focht, and W.J. Farmer. 1987. Evaluation of pesticide groundwater pollution potential from standard indices of soil-chemical adsorption and biodegradation. J. Environ. Qual. 16:422-428.

Karickhoff, S.W. 1984. Organic pollutant sorption in aquatic systems. J. Hydraul. Eng. 110:707-735.

Lyman, W.J., W.F. Reehl, and D.H. Rosenblatt. 1982. Handbook of Chemical Property Estimation Methods: Environmental Behavior of Organic Compounds. McGraw Hill, NY.

Monsanto Agricultural Company. 1988. PHYS: Physical Property Database.

Nash, R.G. 1988. Dissipation from soil, In R. Grover, ed. Environmental Chemistry of Herbicides: Volume I. pp. 131-170, CRC Press, Inc., Boca Raton, FL.

Rao, P.S.C., A.G. Hornsby, and R.E. Jessup. 1985. Indices for ranking the potential for pesticide contamination of groundwater. Proc. Soil Crop Sci. Soc. Fla. 44:1-8.

Wilkerson, M.R., and K.D. Kim. 1986. The Pesticide Contamination Prevention Act: Setting Specific Numerical Values. California Department of Food and Agriculture, Environmental Monitoring and Pest Management, Sacramento, CA.

MATHEMATICAL MODELING OF SEDIMENT LOOP RATINGS: PROBLEM AND LITERATURE REVIEW

Donald R. Jackson[1]

ABSTRACT

In this first of a series of four papers, procedures for collecting water quality data and purposes of data analysis are described. Literature review shows that existing procedures for data analysis are not appropriate to the problem, particularly the description of the loop in concentration versus discharge graphs. Two approaches to the problem of describing the loop in the concentration versus discharge curve are presented.

PROBLEM

The Susquehanna River Basin Commission, in cooperation with USGS, is collecting sediment and nutrient data at 13 sites in the Susquehanna Basin, as part of an ongoing monitoring effort under the EPA Chesapeake Bay Program. The location of the sites are shown in figure 1, and site data is shown in table 1.

The sampling procedure includes collecting data once per month under "baseflow" conditions and also collecting several samples during selected storms. The storms are selected to be representative of different conditions during the year. Most sites are sampled by hand, but automatic samplers have been used for some events at a few sites. Sites having automatic samplers are shown in table 1.

The data is being analyzed for the following purposes:

To compute annual loads;
To compute future trends in loads;
Design sampling procedure to obtain data necessary, but not collect more than the necessary data.

Graphical and statistical analyses of the data were performed using standard techniques. Loads were computed from measured concentration times measured discharge, then all the measurements made under different conditions were lumped together to compute a log-linear regression of load on discharge. As pointed out by McBean and Al-Nassri (1988) this procedure creates a problem with spurious correlation. It also violates assumptions of standard fixed-effects regression analysis. Graphical analysis of data from several different runoff events on the Conestoga River at Conestoga PA shows that different events plot differently as shown in figure 2. This result suggests that several data stratifications should be used in computing the regressions. The regression analysis was repeated using different data stratifications.

While the results were usable, they were not satisfactory because the residual plots were questionable and because some stratifications were not significant.

It was concluded that the standard technique for computing loads by regression analysis is not adequate because of the violation of assumptions of regression analysis.

[1]Donald R. Jackson, Staff Hydrologist, Susquehanna River Basin Commission, Harrisburg, PA

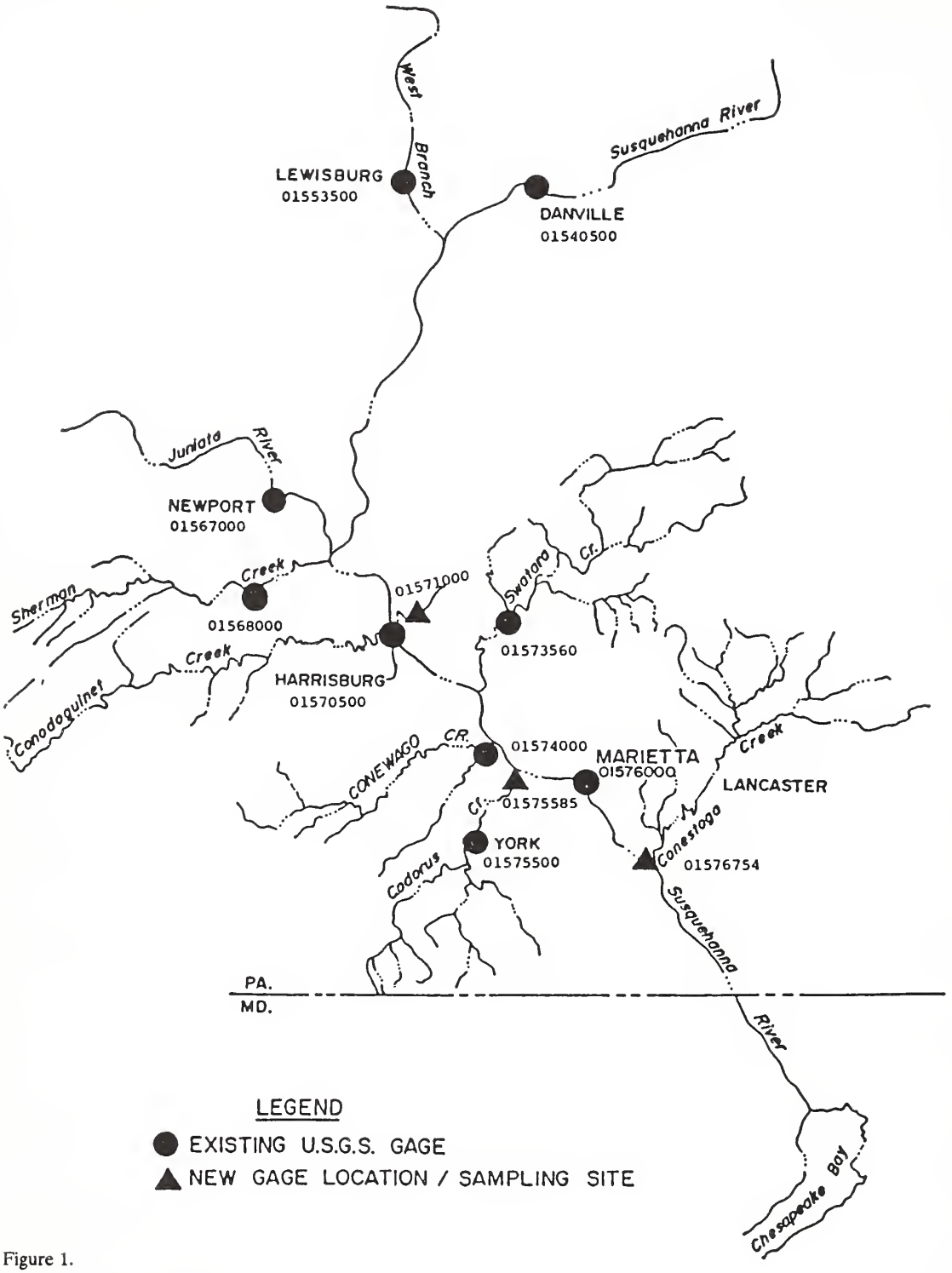


Figure 1.
Sampling site location map.

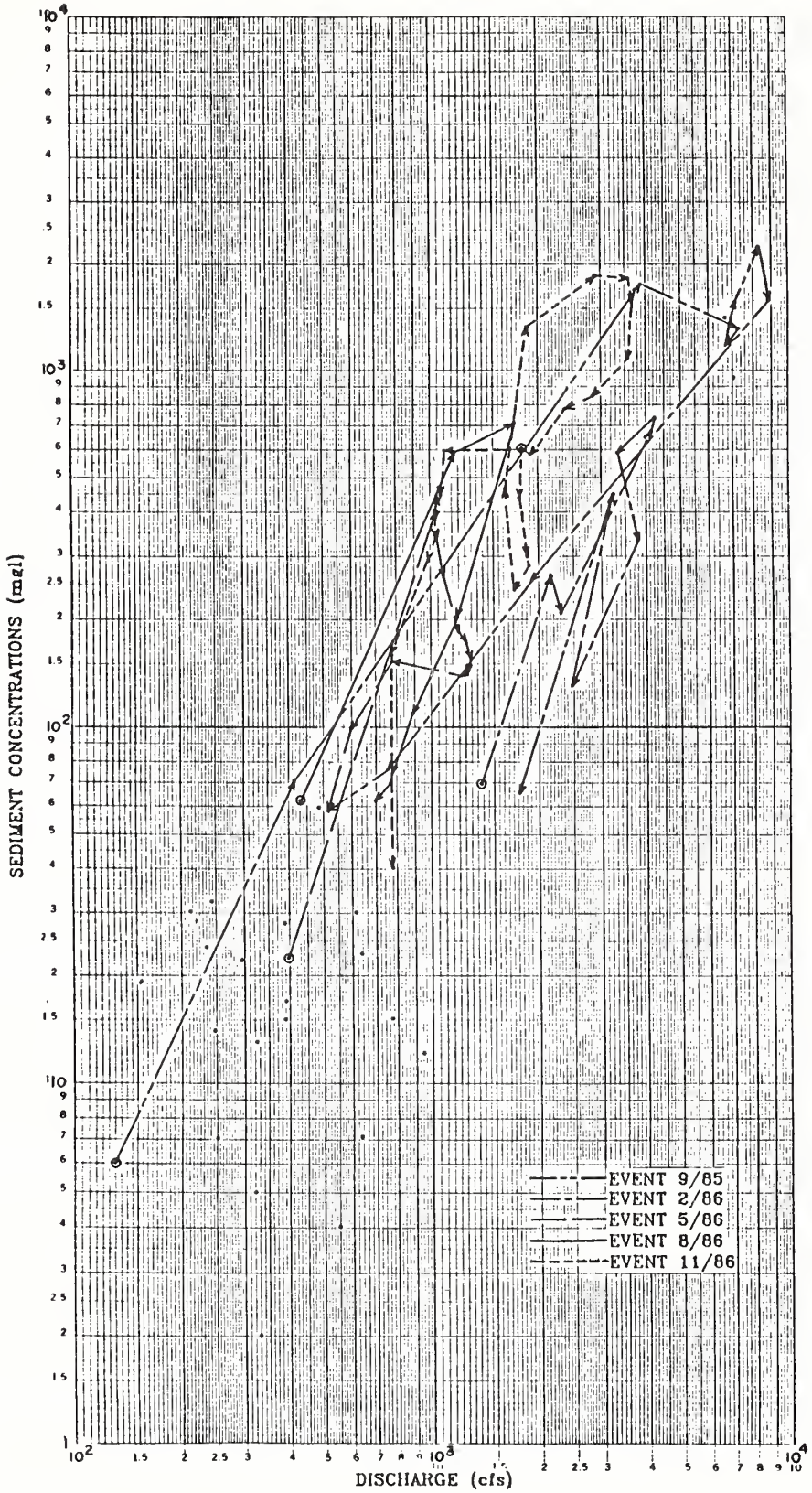


Figure 2.
Loops in sediment concentration versus discharge, Conestoga River at Conestoga.

Table 1.
SRBC/USGS sediment and nutrient monitoring sites.

Sampling Site	Drainage Area Sq. Mi. 2	Period of Record	Number of Storms	Type
Susquehanna R. @ Danville	11,200	Oct 84-Feb 88	6	H
W. Br. Susq. R. @ Lewisburg	6,847	Oct 84-Feb 88	5	H
Juniata R. @ Newport	3,354	Oct 84-Feb 88	5	H
Sherman Cr. @ Shermans Dale	200	Oct 84-Feb 88	7	H
Susquehanna R. @ Harrisburg	24,100	Oct 84-Feb 88	7	H
Paxton Cr. Nr. Penbrook	11.2	Oct 84-Feb 88	12	H
Swatara Cr. Nr. Hershey	483	Oct 84-Feb 88	10	H
W. Conewago Cr. Nr. Manchester	510	Oct 84-Feb 88	11	H
Codorus Cr. Nr. York	222	Oct 84-Feb 88	11	H
Codorus Cr. @ Pleasureville	267	Oct 84-Feb 88	10	A
Susquehanna R. @ Columbia	26,116	Oct 86-Feb 88	2	H
Conestoga R. @ Conestoga	470	Oct 84-Feb 88	13	A

The purpose of the mathematical modeling is to determine whether a model could be developed which would represent the data and facilitate the purposes of the analysis. In particular, it is desirable to develop a model for the purpose of estimating trends in the loads over time, in order to determine compliance with the goals of the Chesapeake Bay Program. The Chesapeake Bay Program is using a base year concept to measure progress towards nutrient reduction goals. However, the inherent variability in hydrology may mask any progress toward meeting the nutrient reduction goal.

The basic concept then is that a model is needed for the relationship between concentration and discharge. The model should include both sediment and nutrients. However, only sediment has been considered thus far.

LITERATURE REVIEW

Review of the literature showed that a model has been developed for the entire Chesapeake Bay drainage which is based on Hydrologic Simulation Program - Fortran (HSPF) (Northern Virginia Planning and Development Comm. 1983). There are many questions regarding this model and it is undergoing major changes.

The questions include:

a. Sensitivity of the model to installation of Best Management Practices.
b. High degree of aggregation.
c. Model doesn't represent the data very well. In particular, it doesn't consider the loops in the sediment rating curves during runoff events.
d. The model needs to be tested against new data presently being collected.

Moore (1984) developed a complex conceptual model which includes:

a. Sediment availability represented by exponential function of time;

b. Sediment removal function which is an exponential decay function of the integral of the direct runoff to an unknown power;
c. A very complex model of runoff generation;
d. Translation model for runoff and sediment reaching the stream.

Moore (1984) applied this model to the River Creedy, U. K. He used continuous data for discharge and sediment. Such data is not available for this problem.

Study of Moore's model shows that the sediment removal function has a maximum value only at the peak of the hydrograph, which ignores the observation that sediment discharge often leads the hydrograph.

Yoo and Molnau (1987) developed a model for sediment generation on small watersheds, which uses the USDAHL model for runoff, and a tractive force concept for the land phase and the channel phase of the sediment generation process. The model appears to be not applicable to large watersheds which are being considered in this study. The results of their model do not look promising for the present purpose.

Rinaldo and Marani (1987) developed a unit response model for runoff and nutrients, which appears to be applicable to sediment also. There are three different formulations of the unit response functions shown. The first is a generalized form of the Nash model for the instantaneous unit hydrograph. The second is a geomorphological instantaneous response function, and the third is a state transition formulation. The relationship between these alternative formulations is not clear. The authors applied the state transition model to a small watershed. The applicability of the state transition formulation to large watersheds seems questionable. The state transition approach seems appropriate for watersheds where rainfall is reasonably uniform spatially.

The conclusion is that there is nothing in the literature which addresses the problem. Because of the type of data, the initial statistical approach, and the desire for simplicity, the writer is attempting to develop model for the sediment loop rating.

Two different approaches to developing such a model have been attempted. The first is an hydraulic model based on the assumption that the loop in the sediment discharge graph is caused by the dynamics of streamflow combined with either changes in sediment transport capacity or with changes in wash load. This model is discussed in the second paper. The second approach fits either a triangular model or a gamma function model of the concentration versus time graph. The second approach is described in the third and fourth papers.

LITERATURE CITED

McBean, E. A., and S. Al-Nassri. 1988. Uncertainty in suspended sediment transport curves. J. Hydraul. Eng. ASCE v. 114, no. 1.

Moore, R. A. 1984. A dynamic model of basin sediment yield. Water Resour. Res. v. 20, no. 1.

Northern Virginia Planning and Development Commission. 1983. Chesapeake Bay basin model. Prepared for US Environmental Protection Agency.

Rinaldo, A., and A. Marani. 1987. Basin scale model of solute transport. Water Resour. Res. v. 23, no. 11, Nov.

Yoo, K. H., and M. Molnau. 1987. Upland soil erosion simulation for agricultural watersheds, Water Resour. Bull. v. 23, no. 5.

MATHEMATICAL MODEL OF SEDIMENT LOOP RATINGS: HYDRAULIC MODEL

Donald R. Jackson[1]

ABSTRACT

In this second of a series of four papers, the development and testing of a hydraulic model to represent the loop in sediment versus discharge graphs is presented. The model is based on equations for the loop in the streamflow rating and a power function for the relationship of concentration versus discharge. The model performs reasonably well on some events, but is considered unsatisfactory in its present form.

CONCEPT

The basic concept of the hydraulic model is that the loop in the sediment concentration versus discharge curve is caused by the dynamics of unsteady channel flow during runoff combined with changes in sediment transport capacity. Henderson (1963, 1966) has shown that the loop in the stage- discharge rating resulting from the dynamics of unsteady channel flow can be represented by the following equations:

$$\frac{Q}{Q_0} = \sqrt{1 + \frac{2}{3r^2} + \frac{5\,Fr^2}{6r^2} + \frac{1}{S_0 c}\frac{\partial y}{\partial t}\left(1 - \frac{Fr^2}{4}\right)} \qquad [1]$$

$$\frac{Q}{Q_0} = \sqrt{1 + \frac{Fr^2}{2r^2} + \frac{1}{S_0 c}\frac{\partial y}{\partial t}\left(1 - \frac{Fr^2}{4}\right)} \qquad [2]$$

where Q = True discharge (cfs) at depth y (ft)
Q_0 = Discharge (cfs) assuming normal flow
r = Ratio of bottom slope to wave slope
Fr = Froude number
S_0 = Bottom slope (ft/ft)
c = Wave celerity (fps)
$\partial y/\partial t$ = Rate of change of depth at a section

Equation 1 applies to crest region of the hydrograph, while equation 2 applies to the flanks.

In order to use Henderson's equations, the celerity computation was modified to represent a natural channel that is not wide.

$$c = \frac{5v}{3} - \frac{4Q}{3T_p}\frac{\partial p}{\partial y} \qquad [3]$$

where v = velocity (fps)
T = Top width (ft)
p = Wetted perimeter (ft)

[1]Donald R. Jackson, Staff Hydrologist, Susquehanna River Basin Commission, Harrisburg, PA

Henderson's equation for defining the crest region was modified as follows:

$$t_p - \frac{2Y_0}{cS_0} < t < t_p + \frac{2Y_0}{cS_0} \quad [4]$$

It was determined that it was necessary to force fit to the USGS rating. Yang's (1973) equation was used to simulate sediment transport.

This model was applied to one event on the Conestoga River at Conestoga PA. The results showed that the model substantially overestimated sediment transport. This result indicates that the sediment concentration on this watershed is controlled by sediment supply or that the wrong sediment transport function was used. Williams and Julien (1987) show that Yang's equation is most appropriate for sand transport. However, the sediment on this watershed is predominantly clay and silt. The conclusion is that sediment supply is the controlling factor, and therefore wash load needs to be modeled.

Current practice is to use a power function to represent wash load (Vanoni 1975, McTernan et al 1987):

$$C = \beta Q^b \quad [5]$$

C = Concentration (mg/l)

In order to represent the loop in the concentration discharge function, the coefficient in equation 5 was made a function of time, leading to the following equation:

$$C = \beta_0 e^{-Kt} Q^b \quad [6]$$

This equation was fitted using least squares in log space:

$$Ln(C) = Ln(\beta_0) - Kt + b\ Ln(Q) \quad [7]$$

This model was fitted to 5 runoff events on the Conestoga watershed. The fitted model and observed data are shown for two events in figures 1 and 2. The regression results are summarized in table 1. The model shows reasonable fit to 3 of the 5 events. The other events suffer from data problems and were not expected to be fitted very well.

The hydraulic model wasn't used in fitting equation 7 because the hydraulic model run for the September 1985 event didn't fit the USGS rating very well and because the loop in the stage-discharge rating for that event was very narrow, indicating little dynamic effect for this watershed and event. The conclusion is that the loop in the concentration versus discharge graph is due to dynamics of wash load supply and not to dynamics of transport in the stream. There is a slight improvement in the explained variance if the hydraulic model is used to compute flows instead of the streamflow rating.

The model is useful for the following purposes:

a. Facilitate understanding of physical processes involved in the observed sediment data.
b. Allow estimation of concentration and total load based only on stage (discharge) for events for which no sampling is done.
c. Allow estimation of concentration and total load for events where insufficient sampling is done.

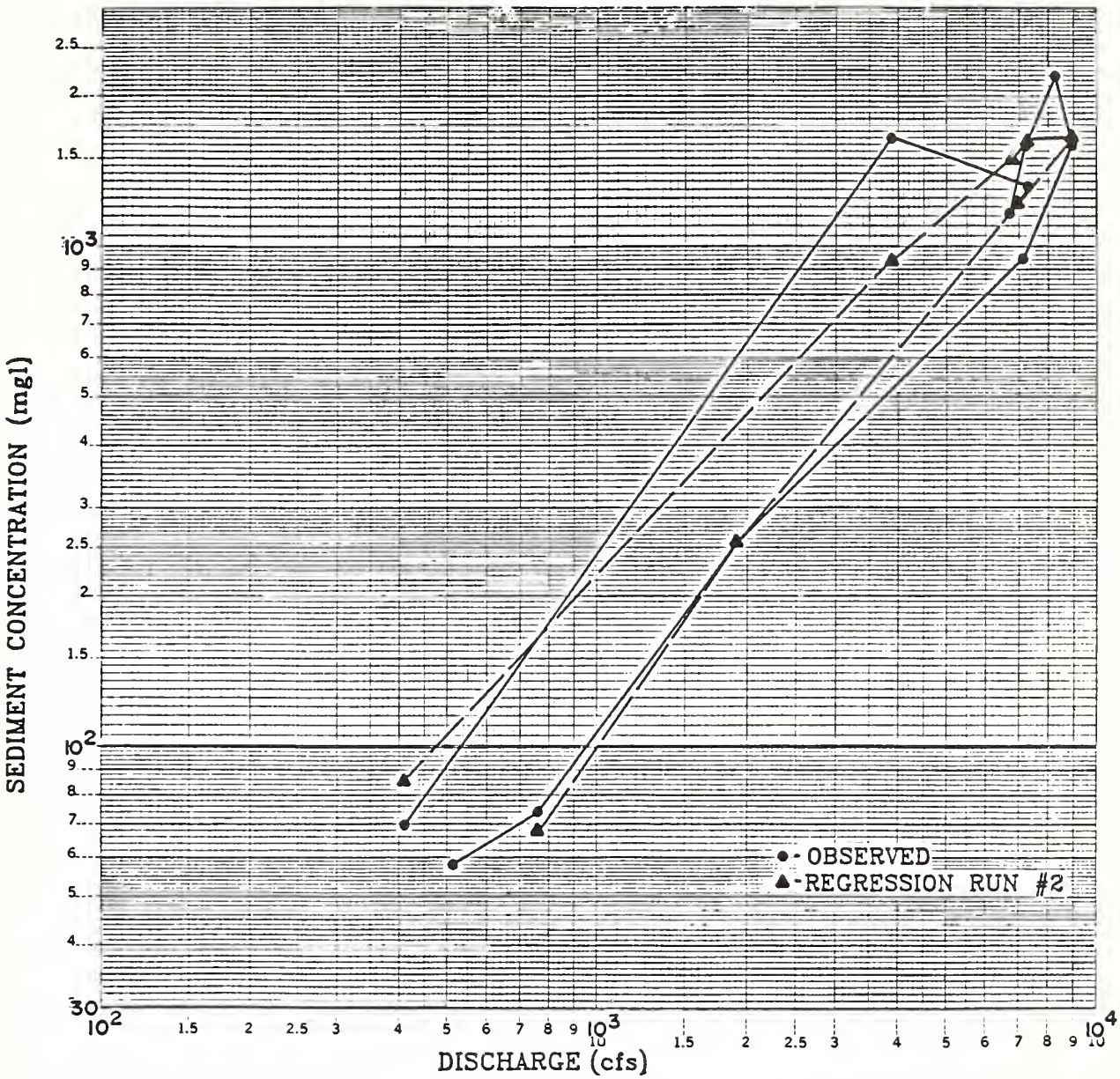


Figure 1.
September 1985 runoff event, Conestoga River at Conestoga.

d. Parameterize process so that factors affecting the process can be studied.
e. May eventually reduce amount of sampling that is needed.

Further work may improve the model; it isn't considered satisfactory in its present form.

Table 1.
Summary of Results of Fitting Sediment Model.

Event Date	Intercept	Coefficient For Time	Coefficient For Ln(Q)	R^2	Error Degrees of Freedom
9/85	- 2.1350	-0.01625	1.1075	0.95	7
2/86	-10.6638	-0.004734	2.0968	0.90	7
5/86	- 5.8141	0.003980	1.5568	0.59	10
8/86	- 8.248	-0.03153	2.0391	0.91	4
11/86	- 4.829	-0.006689	1.4938	0.74	16

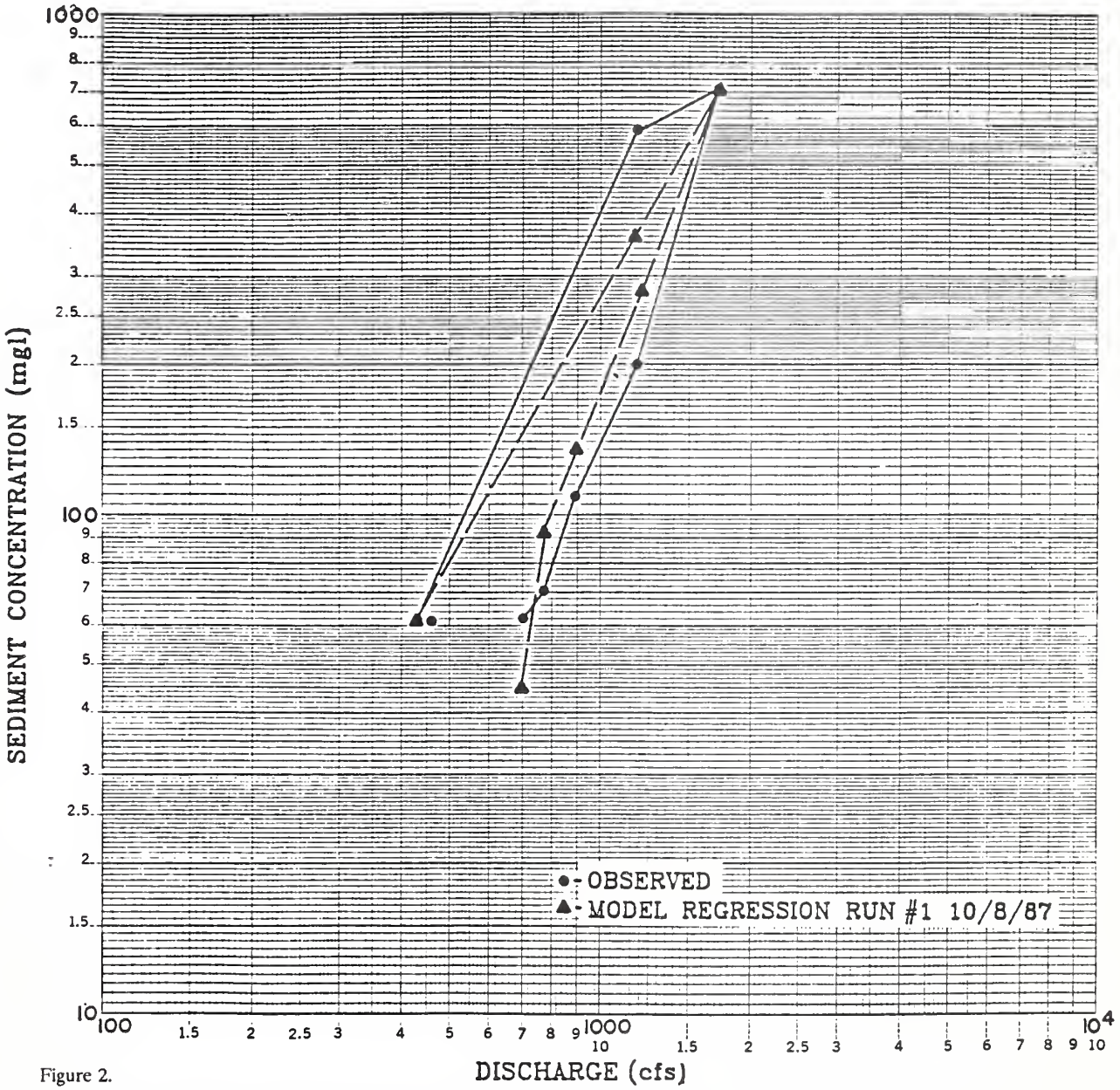


Figure 2.
August 1986 runoff event, Conestoga River at Conestoga.

LITERATURE CITED

Henderson, F. M. 1963. Flood waves in prismatic channels. J. Hydraul. Div. ASCE v. 89, no. HY4.

Henderson, F. M. 1966. Open channel flow. Macmillan Co. pp. 374-394

McTernan, W. F., B. L. Weand, and T. J. Grizzard. 1987. Evaluation of management practices to control agricultural pollutants. Water Resour. Bull. v. 23, no.4.

Williams, D. T., and P. Y. Julien. 1987. On the selection of sediment transport equations. Proceedings Conference on Hydraulic Engineering ASCE.

Vanoni, V. ed. 1975. Sedimentation engineering. ASCE pp. 475-7.

Yang, C. T. 1973. Incipient motion and sediment transport. J. Hydraul. Div. ASCE v. 99, no. HY10.

MATHEMATICAL MODELING OF SEDIMENT LOOP RATINGS: TRIANGULAR AND GAMMA FUNCTION MODEL DEVELOPMENT

Donald R. Jackson[1]

ABSTRACT

In this third in a series of four papers, the development of triangular and gamma function models to represent sediment concentration versus time graphs is presented.

CONCEPTS

The studies described in the previous papers in this series led to the question: What creates the loop in the concentration versus discharge relationship. The obvious (albeit superficial) answer is that the loop is due to the sediment concentration leading the discharge. This line of thinking suggested that the time relationship between the hydrograph and the pollutograph should be modeled. Some investigators have used triangles to represent hydrographs; others have used a continuous function such as the gamma probability density function (Reich 1962). Thus these two functions were studied to see if they could also represent the pollutograph, and the loop in the concentration versus discharge graph.

PROCEDURE

Initial computer simulations used triangles to understand the relationship between the pollutograph and hydrograph, and to verify the assumption that the lead-lag relationship between the two creates the loop in the concentration versus discharge graph. Then computer programs were written to fit the triangle and the gamma function to the observed pollutograph. It was assumed in modeling actual data that the discharge hydrograph was known from the measured stage hydrograph and the rating curve, and thus it isn't necessary to model the hydrograph.

EXPERIMENTS WITH TRIANGULAR MODEL

The relationships between the triangular representation of the runoff hydrograph and the pollutograph are shown in figure 1. These relationships are defined by:

$$C_p = \xi \quad Q_p \qquad T_b = \pi_4 \; T_p$$

$$\tau_p = \pi_2 \; T_p \qquad \tau_b = \pi_3 \; T_p$$

The equations for the triangles become:

$$Q(t) = \frac{t}{T_p} \; Q_p \qquad 0 \leq t \leq T_p$$

$$Q(t) = \frac{Q_p(t-T_b)}{T_p - T_b} \qquad T_p \leq t \leq T_b$$

[1]Donald R. Jackson, Staff Hydrologist, Susquehanna River Basin Commission, Harrisburg, PA

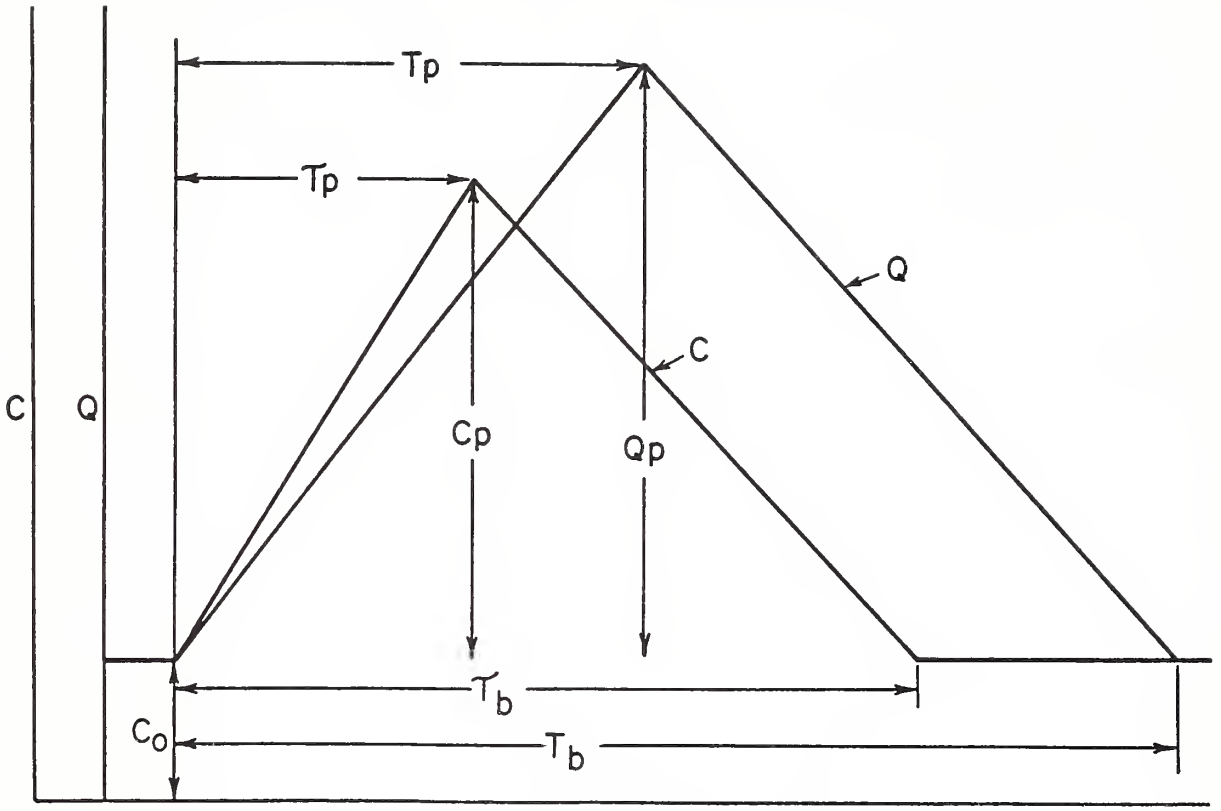


Figure 1.
Definition sketch for triangular models for concentration and discharge versus time.

$$C(t) = \frac{\xi\, t\, Q_p}{\pi_2\, T_p} \qquad 0 \leq t \leq \pi_2\, T_p$$

$$C(t) = \frac{\xi\, Q_p(t-\pi_3\, T_p)}{\pi_2\, T_p - \pi_3\, T_p} \qquad 2T_p < t < 3T$$

Simulations were made with the following values of the parameters:

Q_p	=	0.05, 0.25, 0.5 in/hr
T_p	=	4.0, 10.0, 16.0 hr
ξ	=	1.0
π_4	=	2.0, 2.5, 3.0
π_2	=	0.6, 0.8, 1.0
π_3	=	1.25, 1.50, 1.75

Discharge and concentration versus time data were simulated for all (243) possible combinations of parameters. Height and width of the loop, and volume within the loop were computed. Discharge and concentration versus time and log(C) versus log(Q) plots were prepared. Typical results are shown in figures 2 and 3. Plots of dependent variables versus the parameters of the model were also generated but not included here. Mathematical and empirical studies of properties of system of triangles were performed but are not included here.

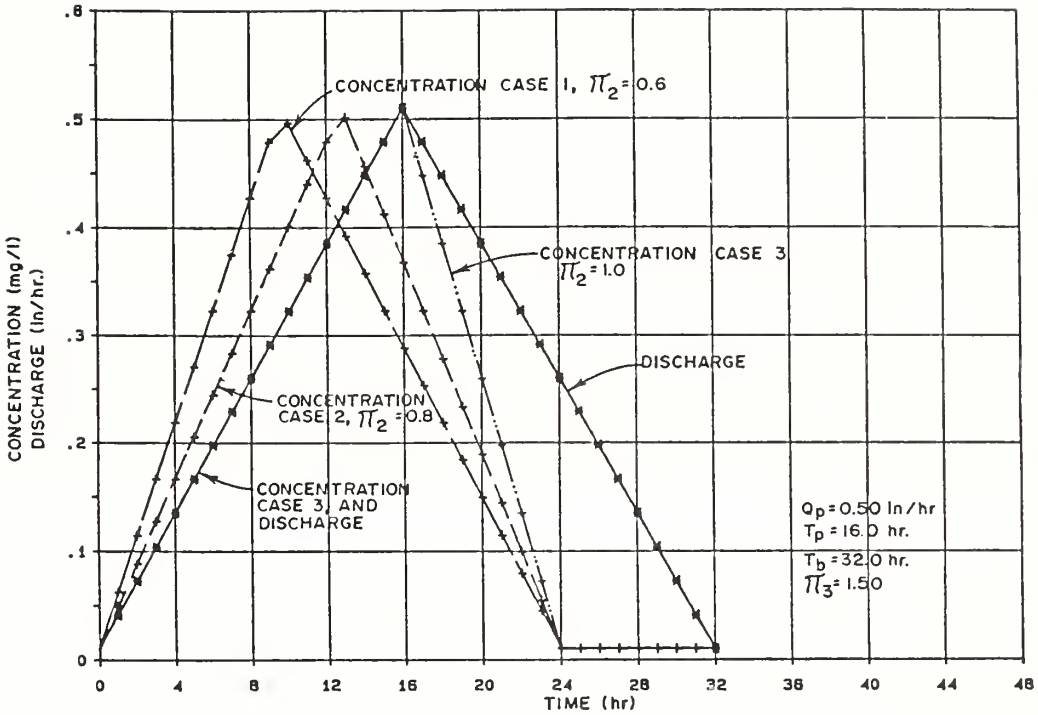


Figure 2.
Typical plot of concentration and discharge versus time.

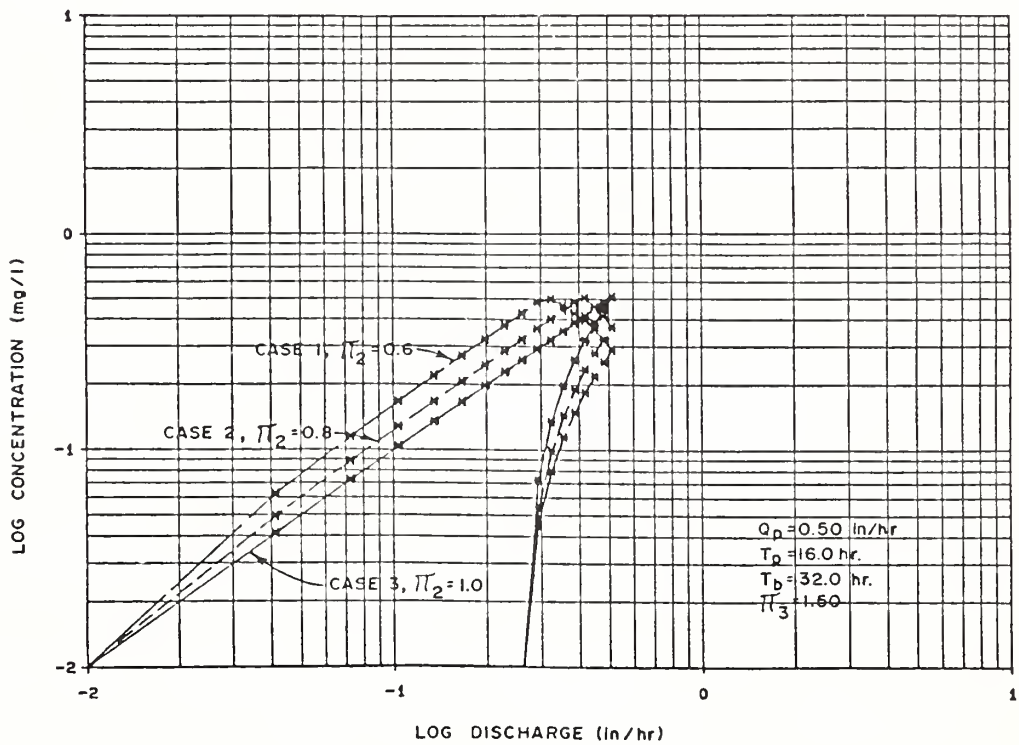


Figure 3.
Typical plot of log (concentration) versus log (Discharge).

The conclusions are:

1. Triangular representation for both hydrograph and pollutograph may not be valid, because of tendency for C-Q loop not to close.
2. Inverted V shape given by power equation with exponential decay for C vs. Q , used to model wash load in the hydraulic model, occurs only when $\tau_p = T_p$. Thus power equation doesn't represent the case where sediment concentration leads the discharge.

MATHEMATICAL REPRESENTATION OF TRIANGULAR MODEL OF POLLUTOGRAPH

Fitting the triangle to observed data reduces to fitting a function of the form

$$C(t) = C(t_0) + b(t-t_0)$$

separately to the rising and falling parts of the pollutograph. The fitting procedure considers two cases. In the first case, there are only two points on the rise or recession of the pollutograph. In this case, the slope and intercept of the straight line is uniquely determined by the two points. In the second case, there are more that two data point on either the rise or the recession. In this case, least squares was used to fit the respective points. The peak concentration and time to peak was determined by intersection of the two straight lines. The time of beginning of rise was determined by projecting the line to an assumed initial concentration. The end of the recession was determined by projecting the fitted line to an assumed final concentration. The goodness of fit for the second case was determined by applying the standard equations for the R^2 and F statistics to the rise and recession separately. Overall goodness of fit was determined for both cases by computing an equivalent R^2 and F statistic based on computed residuals. Degrees of freedom for the F statistic were assumed to be 2 and N-4, where N is the number of data points available.

MATHEMATICAL REPRESENTATION OF GAMMA FUNCTION MODEL OF POLLUTOGRAPH

The gamma function is represented by (Rinaldo and Marani, 1987; Reich 1962, Freund 1962):

$$C(t) = C(t_0) + \frac{F}{K\Gamma(\alpha)} \left(\frac{t}{K} \right)^{\alpha-1} e^{-t/K} \qquad [1]$$

where $C(t)$ = concentration as function of time t (mg/l)
$C(t_0)$ = concentration at beginning of rise (mg/l)
K, α, are parameters
F = scaling factor equal to the area

The initial assumption was that the function could be fitted using either least squares in log space, or by the method of moments.

The least squares solution can be found by taking logarithms of equation 1:

$$\text{Ln}\,[C(t)-C(t_0)] = \text{Ln} \left(\frac{F}{K^{\alpha}\Gamma(\alpha)} \right) + (\alpha-1)\,\text{Ln}\,(t) - \frac{t}{K} \qquad [2]$$

which has the form

$$Y = B_0 + B_1 \, Ln(t) + B_2 \, t \qquad [3]$$

$$B_0 = Ln \left(\frac{F}{K^{\alpha} \Gamma \alpha} \right) \qquad [4]$$

$$B_1 = \alpha - 1 \qquad [5]$$

$$B_2 = - \frac{1}{K} \qquad [6]$$

$$F = \left(- \frac{1}{B_2} \right)^{B_1 + 1} (B_1 + 1) \exp(B_0) \qquad [7]$$

The method of moments solution (Freund, 1982) was developed by assuming that the observed data points were approximated by trapezoids with area A_i and moment arm T_i. Then the first moment about the origin is given by

$$\mu_1 = \bar{x} = FK\alpha \qquad [8]$$

and the second moment about the mean is given by

$$\mu_2 = FK^2\alpha^2(1-F) + FK^2 = s^2 \qquad [9]$$

Then

$$F = \int_0^{\infty} C(t)dt = \Sigma A_i \qquad [10]$$

$$\mu_1 = \frac{\Sigma T_i A_i}{\Sigma A_i} \qquad [11]$$

$$\mu_2 = \frac{\Sigma (T_i - \bar{T})^2 A_i}{\Sigma A_i} \qquad [12]$$

For the least squares solution, R^2 and F statistics were computed using standard equations from statistics for the log linear form. For both cases, equivalent R^2 and F statistics in concentration units were computed by using the standard equations with the residuals computed from the untransformed fitted equation.

LITERATURE CITED

Freund, J. E. 1962. Mathematical statistics. Prentice Hall. pp. 146-7.

Reich, B. M. 1962. Design hydrographs for very small watersheds from rainfall. Civil Engineering Section Colorado State University. CER62BMR41.

Rinaldo, A., and A. Marani. 1987. Basin scale model of solute transport. Water Resour. Res. v. 23, no. 11.

MATHEMATICAL MODELING OF SEDIMENT LOOP RATINGS: TRIANGULAR AND GAMMA FUNCTION MODEL RESULTS

Donald R. Jackson[1]

ABSTRACT

In this fourth of a series of four papers, the results obtained from fitting triangular and gamma function models to the sediment concentration versus time graph are described. The models seem to be promising but more work is needed.

PRACTICAL PROBLEMS

The following problems were encountered in fitting the triangular and gamma function models.

a. For some events, the initial concentration may be fairly apparent, in other cases it may be a guess.

b. Time of beginning of pollutograph rise has to be selected with some care, especially for the triangular model.

c. Points on the recession tail or at the beginning of the pollutograph may need to be ignored in fitting either the triangular model or the gamma model.

d. Certain events cannot be analyzed using the triangular model at this time. Double peaked events are difficult for both models, but especially the triangular model.

e. For both models, location of peak is unknown. This is particularly a problem in fitting the triangular model because maximum observed concentration may be part of the rise or the recession or both.

RESULTS OF TRIANGULAR MODEL

The triangular model was fitted to four selected events on Conestoga River at Conestoga, PA. Goodness of fit statistics are shown in table 1. Graphs for two of the events are shown in figures 1 and 2.

Results of fitting the triangular are as follows.

The September 1986 event is double peaked. Consequently, the slope of rise is too flat, and the equation for the rise is not significant. The recession seems to be fitted well. The overall F statistic is significant at 5% (one sided) but not at 1%. Sensitivity analysis of time of beginning of rise shows that the equation for the rise remains the same. Sensitivity analysis of the assumption regarding location of peak shows that including the peak value in computing the equation of the rise increases the slope of the rise, and shifts the intercept to later time. Changing the assumed location of peak also results in a more consistent model. The intercept for the initial concentration is too early regardless of assumed location of peak.

[1]Donald R. Jackson, Staff Hydrologist, Susquehanna River Basin Commission, Harrisburg, PA

Table 1.
Summary of computer output triangular model Conestoga River at Conestoga.

Event Date	Soln. No.	Goodness of Fit Statistics									Tp (hr)	Cp Comp (mg/l)	Cp Obs. (mg/l)	Cp Comp % of Obs.
		Rise			Recession			Overall						
		RSQ	F Statistic Value	F Statistic DF	RSQ	F Statistic Value	F Statistic DF	RSQ	F Statistic Value	F Statistic DF				
9/85	1	0.553	4.94	1/4	0.967	57.89	1/2	0.79	13.21	2/7	28.45	2071	2220	93
9/85	2	0.342	1.56	1/3	0.967	57.89	1/2	0.80	13.95	2/7	29.43	1939	2220	87
9/85	3	0.553	4.94	1/4	0.967	57.89	1/2	0.79	13.21	2/7	25.45	2071	2220	93
9/85	4	0.342	1.56	1/3	0.967	57.89	1/2	0.80	13.95	2/7	26.43	1939	2220	87
8/86	1	1.00	---	---	1.00	---	---	0.996	366	2/3	11.94	781	710	110
	2	1.00	---	---	1.00	---	---	0.996	229	2/2	11.94	781	710	110
1/87	1	0.946	35.30	1/2	0.947	17.81	1/1	0.920	11.58	2/2	6.29	713	733	97
3/87	1	0.922	47.12	1/4	0.969	124.4	1/4	0.930	92.61	2/14	24.18	1182	1198	99

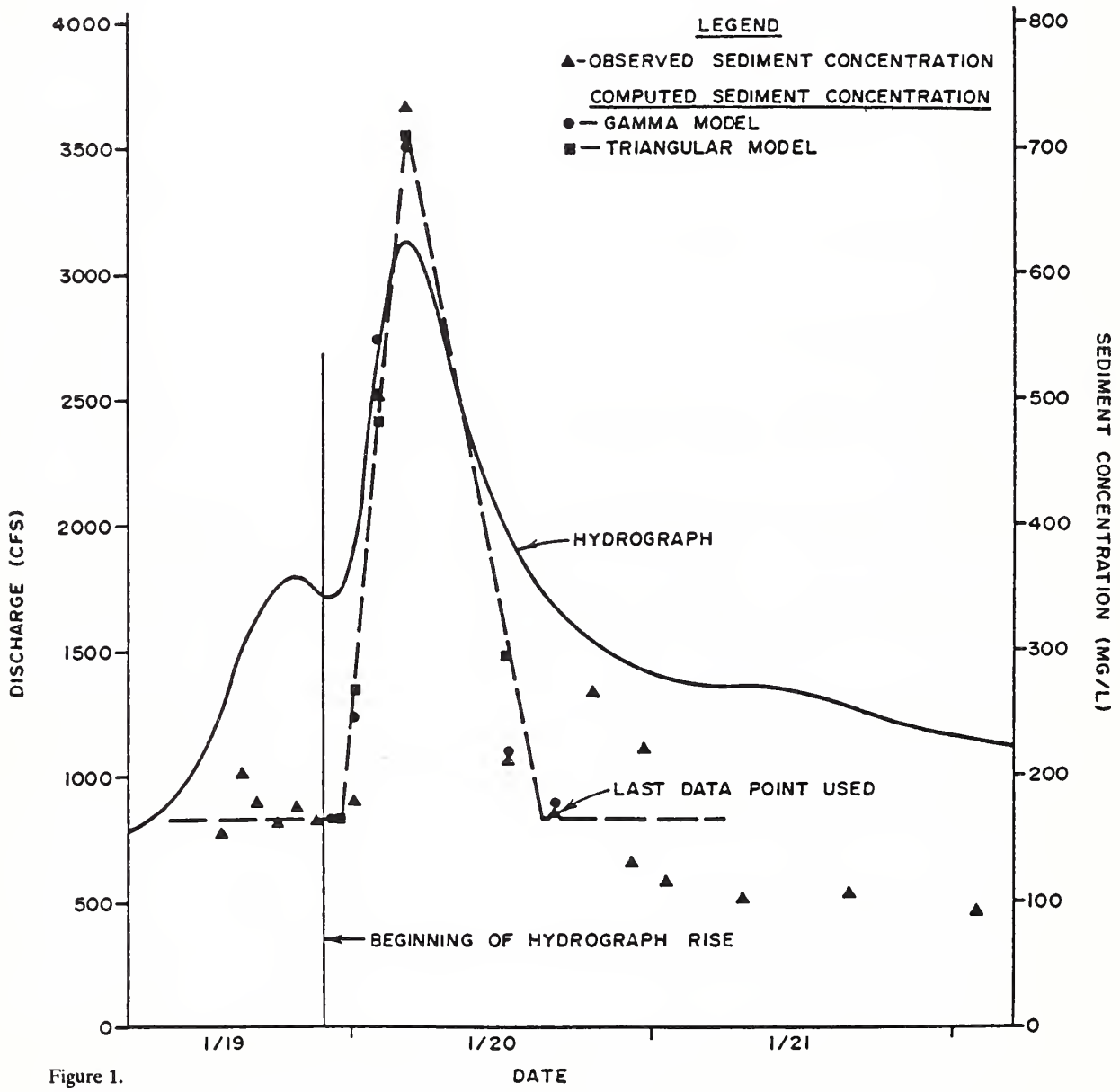


Figure 1.
January 1987 Event, Conestoga River at Conestoga.

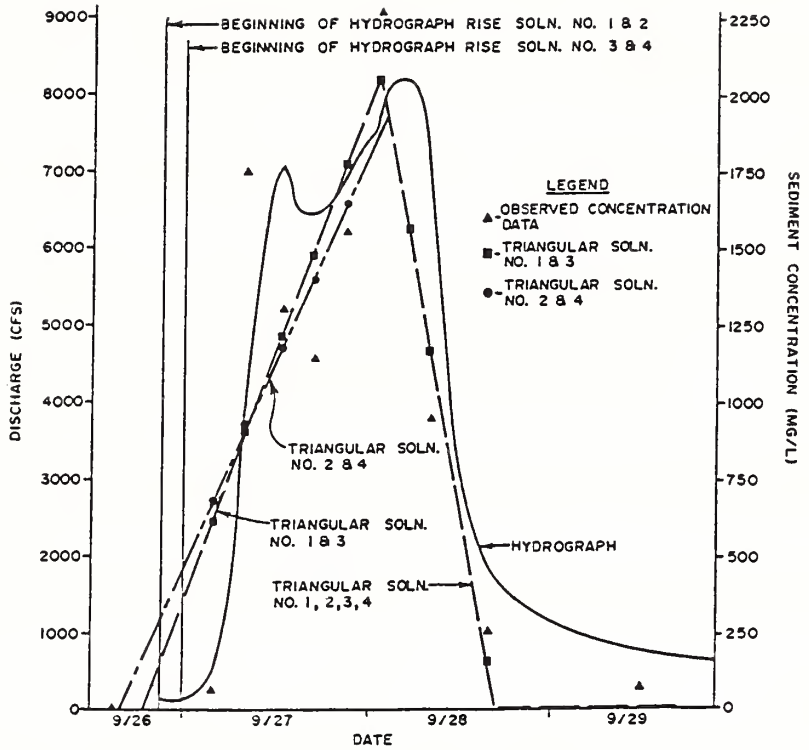


Figure 2(a).
September 1985 event, Conestoga River at Conestoga.

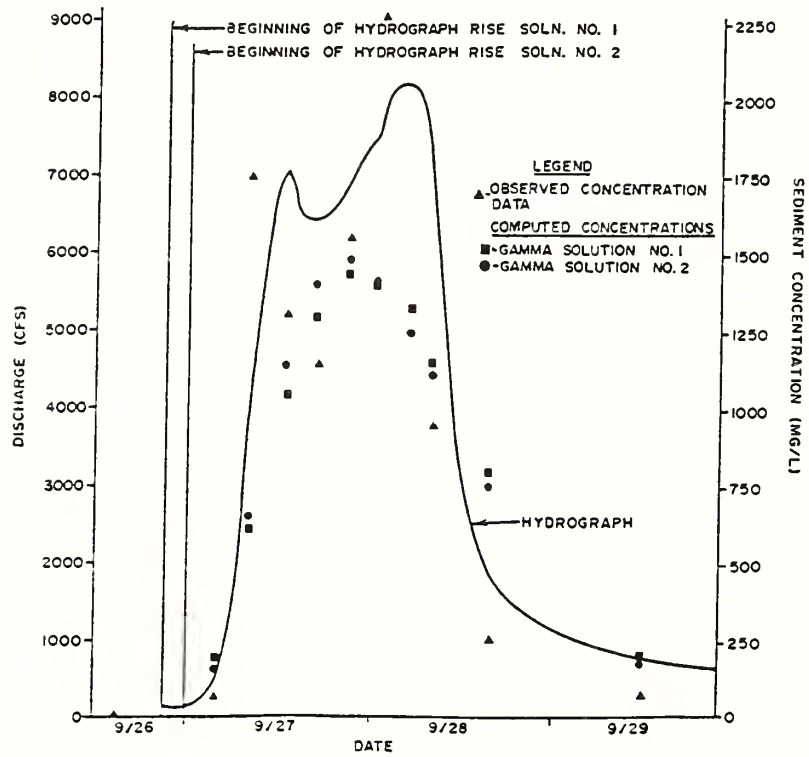


Figure 2(b).
September 1985 event, Conestoga River at Conestoga.

The August 1986 event has two points on the rise and two on recession. The fit is very good, as it should be.

The January 1987 event is fitted satisfactorily. There is some question regarding time and concentration at the beginning of the rise, based on the hydrograph.

The March 1987 event is fitted satisfactorily. There is some question regarding initial concentration, but changing the initial concentration wouldn't have significant effect on goodness of fit.

RESULTS OF GAMMA MODEL

The gamma function was fitted to the same four events as the triangular model. Typical results are shown in figures 1 and 2. The goodness of fit is summarized in table 2.

It was found that method of moments doesn't work. One reason is that the value of K is determined as the small difference between relatively large numbers.

For the least squares solution the intercept is sometimes not significantly different from zero, which makes it difficult to compute the scaling factor F. Instead of computing F from the intercept given by the least squares solution, the value of F was computed from the area under the observed points, using the method of moments solution.

The two independent variables in the least squares solution are highly correlated. It is questionable whether ordinary least squares can be used to fit the gamma function.

The following is a summary of the results of using the least squares solution for individual events.

1. For the Sept. 1985 event, the rise and recession are fitted reasonably well but the peak is significantly underestimated. The F statistic is significant at 1% (one sided) in log units, 5% in concentration units. Sensitivity analysis shows that changing starting time improves goodness of fit slightly. Deleting points from the tail of the recession may improve goodness of fit also.

2. For the Aug. 1986 event, the rise and recession are fitted reasonably well, but the peak is underestimated for the first case. Time to peak is reasonable but perhaps slightly early. F statistic is significant at 1% (one sided) in log units and in concentration units. Sensitivity analysis (not shown on graph) shows that deleting the last point on the recession shifts time of peak to be much too early, but improves goodness of fit in log units (but not in concentration units) and results in overestimating the peak slightly.

3. The January 1987 event is fitted satisfactorily. The F statistic is significant at 5% in log space, 1% in concentration units.

4. For the March 1987 event, the rise and recession are fitted reasonably well, but the peak is underestimated. The F statistic is significant at 1% in both log and concentration units. A better estimate of initial concentration may improve the fit.

5. For the latter three events, where the hydrograph and the pollutograph are single peaked, the time to peak seems reasonably good. Since $T_p = K(\alpha-1)$, the results suggest that the product of the parameters is being estimated well, but that the individual parameter values may not be estimated adequately.

Table 2.
Summary of computer output gamma model, Conestoga River at Conestoga.

Even Date	Soln No.	Goodness of Fit Statistics: Log Units RSQ	Log Units F Statistic Value	Log Units F Statistic DF	Concentration Units RSQ	Concentration Units F Statistic Value	Concentration Units F Statistic DF	Tp (hr)	Cp Comp (mg/l)	Cp Obs. (mg/l)	Cp Comp % of Obs.
9/85	1	0.745	11.7	2/8	0.611	6.27	2/8	25.09	1445	2220	65
9/85	2	0.787	14.8	2/8	0.636	6.99	2/8	20.75	1490	2220	67
8/86	1	0.909	19.88	2/4	0.888	15.81	2/4	10.42	597	710	84
	2	0.988	128.21	2/3	0.859	9.12	2/3	8.98	780	710	110
1/87	1	0.904	14.06	2/3	0.977	64.17	2/3	6.23	700	733.1	95
3/87	1	0.936	101.94	2/14	0.80	27.25	2/14	24.80	828	1198.9	69

CONCLUSIONS

Both the triangular model and the gamma function model perform reasonably well for these events.

The triangular model fits peak reasonably well, but doesn't fit the tail of the recession very well. The gamma function fits the tail of the recession better, but not the peak.

It is probably necessary to develop separate model (e. g. exponential decay) for the recession after some time, especially for the triangular model.

It is not possible to choose between models at this time.

A better fitting algorithm is needed for gamma model.

Concentration versus discharge graphs should be studied.

Procedures for modeling complex multi-peak events need to be developed.

A MODEL OF WATER AND SOLUTE MOVEMENT IN STRUCTURED SOILS

N.J. Jarvis[1]

ABSTRACT

Although the important role of macropores in controlling water and solute movement in structured field soils is now widely recognized (Thomas and Phillips 1979) our level of quantitative understanding is still poor. For example, it is not possible with present knowledge to predict with a reasonable degree of confidence the likely effects of soil management practices (e.g. tillage) on the movement of solute to drains, rivers and groundwater. The variety of factors which affect the macropore system and the complexity of their interaction suggests that a modeling approach offers the best means to improve our understanding of these processes.

A number of models of solute transport applicable to structured soils have been developed. Analytical approaches (van Genuchten and Wierenga 1976) which assume steady-state conditions are not well suited to field applications in which intermittent inputs occur at the soil surface. Addiscott (1984) described a numerical model to predict leaching in structured field soils. In this approach, a simple water balance accounting procedure was used to predict the flow rate of 'mobile' water, and exchange of solute with 'stagnant' water occurred by intra-aggregate diffusion in cubic-shaped peds.

The model briefly described here adopts a description of solute transport processes which is similar to that of Addiscott, but instead this is coupled to a more rigorous physically-based water balance model (Jarvis and Leeds-Harrison 1987a,b). The full model is described in greater detail by Jarvis (1989).

DESCRIPTION OF THE MODEL

Soil Structure

The soil profile is divided into discrete layers, each of which is assumed to contain cubic soil peds of equal size, separated by planar cracks. Total crack porosity is a function of the soil-water deficit and the shrinkage characteristic. Crack width and total ped surface area per unit soil volume are calculated from the crack porosity and crack spacing (= ped size).

Recharge

Philip's infiltration equation is used to partition rain or irrigation into water entering soil peds and water flowing into cracks. Two factors control this process: the rainfall intensity and the ped sorptivity at the soil surface. The latter is assumed a linear function of the soil-water deficit.

Evaporation and Root Water Uptake

Potential transpiration is calculated using a generalized form of Penman's equation in which canopy and aerodynamic resistances are explicitly accounted for.

[1] N.J. Jarvis, Assistant Professor, Department of Soil Sciences, Swedish University of Agricultural Sciences, Box 7014, 750 07 Uppsala, Sweden.

Actual transpiration and root water uptake are estimated as a function of the root depth and distribution and the soil water content.

Water in Peds

Flow in peds is assumed negligible: peds simply act as swelling absorbers and shrinking desorbers of water. Thus, the model is best suited to fine-textured soils.

Water stored in the cracks is taken up into the peds at a rate which depends on the ped sorptivity and the wetted contact area. The latter is calculated knowing the degree of saturation in the cracks and the total ped surface area.

Water in Cracks

Water in the cracks is either mobile (above the wetting front during rainfall) or stagnant ('old water' remaining from previous raindays). During rainfall or irrigation, stagnant water is incorporated into the mobile phase as the wetting front in the cracks advances down the profile.

The degree of saturation in the mobile phase depends on the balance between the input rate at the soil surface, the uptake rate into peds and the flow rate in the cracks.

The flow rate of mobile water is a function of the crack width, crack porosity and an empirical factor related to flow path tortuosity and connectivity.

Drainage

The drains are assumed to respond when the cracks become saturated above drain depth. Two zones of saturation may exist within the soil profile: a continuously fluctuating 'groundwater table' and a transient perched water table.

Seepage potential theory is used to calculate the drainage rate as a function of saturated hydraulic conductivity, the height of the water table above drain depth and the drain spacing.

Runoff

'Saturation excess' runoff occurs if the groundwater table in the cracks reaches the soil surface. 'Infiltration excess' surface runoff occurs if the rainfall intensity exceeds the combined infiltration capacity of peds and cracks: a perched water table or 'infiltration throttle' then develops.

Solute Transport

The peds are divided into segments of equal volume. Fick's law is used to calculate diffusion of solute between segments from the diffusion coefficient in free water, the ped water content, the impedance factor, the mean cross-sectional area and the solute concentration gradient. Solute diffusion between the outermost ped segment and the water-filled cracks is calculated as above, but is also dependent on the degree of saturation in the cracks.

Mass transport of solute occurs in the mobile water in the cracks during rainfall, and also in saturated flow to drains. Solute is also taken up into the outermost ped segment by mass flow (in water absorbed by peds).

The user may allocate different solute concentrations to rainfall and irrigation.

APPLICATIONS

Figure 1 shows an initial test of the water balance model for a heavy clay soil at Silsoe (U.K.) under grass. Soil water contents measured by the neutron probe are compared with model predictions for a 7-week period from 17th April 1984. Figure 1a depicts an initial dry period characterized by soil water depletion to 60 cm depth. Subsequently, the profile was re-wetted by 100 mm of rain (fig.1b). In both cases, the agreement between model predictions and measurements is excellent. It should be noted that only a minimum of model calibration was performed (adjustment of the assumed root distribution), since independent measurements of many model parameter values are available for this site (Jarvis and Leeds-Harrison 1987b).

A simulation of solute leaching to drains in the same soil for a 1-month period in autumn 1984 is summarized in figures 2,3 & 4. Figure 2 shows the initial solute distribution in the soil following application of 50 kg ha^{-1} solute in 50 mm of irrigation water. The pattern predicted by the model is characteristic of structured soils, in that the solute has penetrated to the base of the profile, carried by rapid water flow in cracks. Although much of the applied solute is stored close to the soil surface, a secondary peak or bulge of solute is also observed, in this case at 60-80 cm depth. This is due to the response of the water table in the cracks, leading to an increase in the wetted ped surface area and thus solute uptake (both by mass flow and by diffusion).

Figure 2 also shows the predicted solute distribution after nearly one month of leaching (total rainfall of 53 mm). Most of the solute leached from the profile has been removed from the surface 15 cm. These layers have a finer structure (at least under grass) and are also wetted more frequently. Therefore, solute diffusion to ped surfaces and subsequent leaching is a more efficient process. Some of this leached solute is redistributed to deeper soil layers (20-50 cm depth), but most is lost to the drains.

Figure 2 suggests that solute leaching may be strongly affected if the topsoil structure is changed by management practices (i.e. tillage). As an exploratory example, figures 3 and 4 show the likely effects of a 50% increase and a 50% decrease in the size of peds in the topsoil (0-15 cm).

Figure 3 shows that the first breakthrough to drains was predicted after 10 days in the fine-structured soil, during which time only 16 mm of rain had fallen. Predicted leaching from the medium and coarse-structured soils started 2 days later, after a further 13 mm of rain. After nearly one month, the total solute loss to the drains was still 20% higher in the fine structure than in the coarse. This was the case despite the fact that the predicted total drainflow was 33% lower in the fine-structured soil.

The reason for this is illustrated in figure 4 which shows that the predicted concentration of solute in the water draining from the fine-structured soil is nearly double that found in both other soils.

FUTURE WORK

The main priority is to test the coupled water balance and solute transport model in the field under a variety of conditions. Work in this direction is in progress.

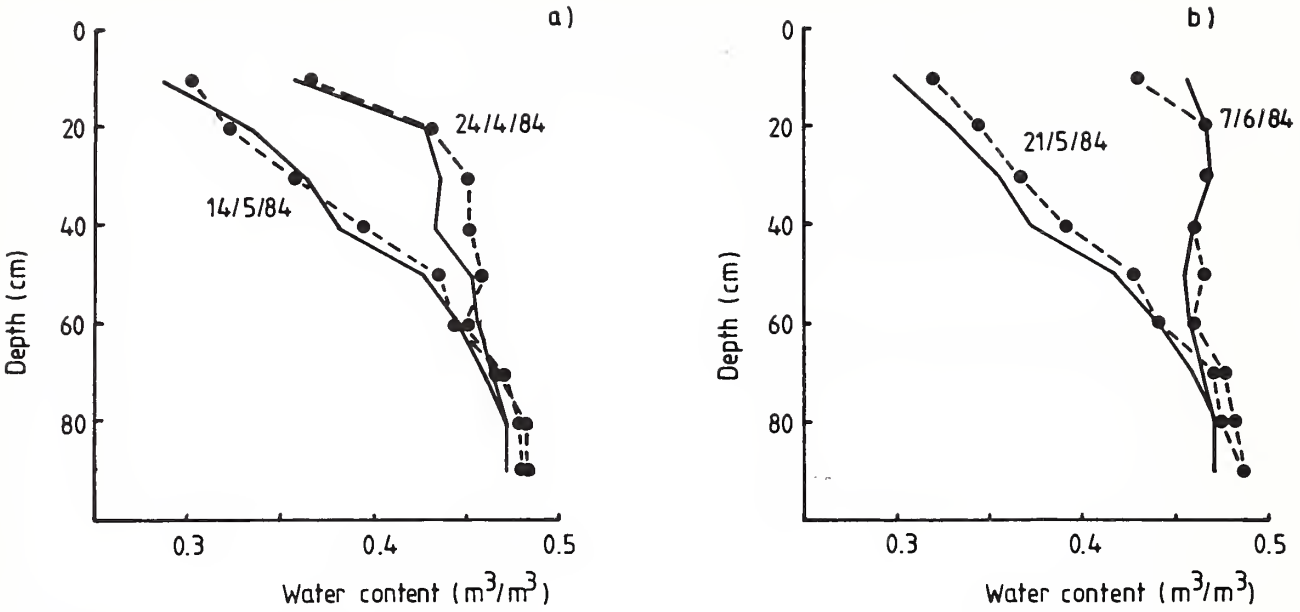


Figure 1.
A comparison of measured (--•--) and predicted (———) soil water content profiles, Evesham clay soil, Silsoe, U.K.

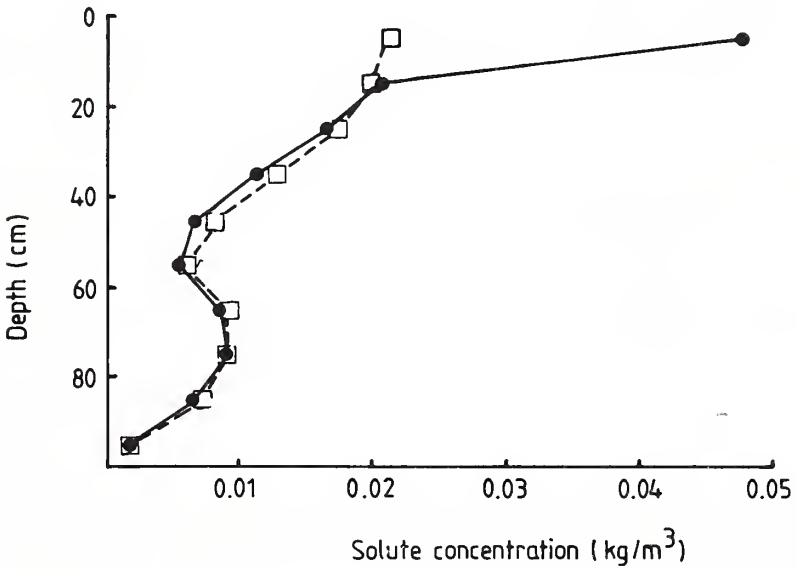


Figure 2.
Solute concentration profiles predicted on day 1 (•) and day 30 (•).

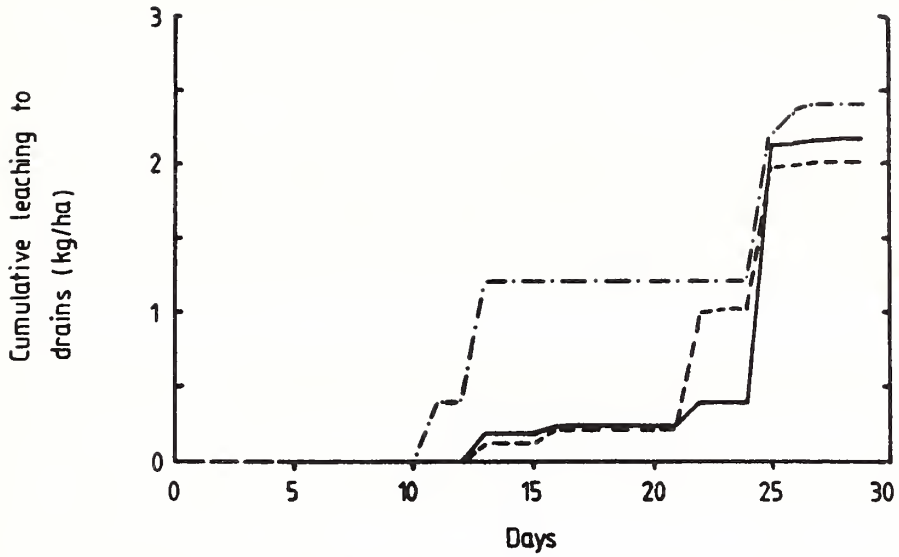


Figure 3.
Cumulative loss of solute predicted in drain outflow (—•—• fine, ——— medium, ------ coarse aggregates).

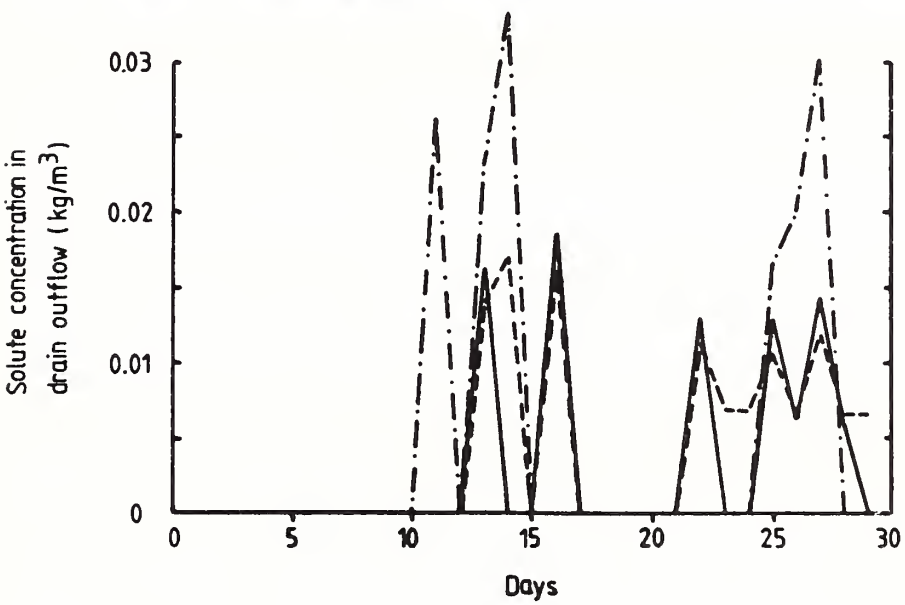


Figure 4.
Solute concentration predicted in drain outflow (—•—• fine, ——— medium,------ coarse aggregates).

ACKNOWLEDGMENTS

The financial support of NFR (Swedish Natural Sciences Research Council), SJFR (Swedish Council for Forestry and Agricultural Research) and AFRC (Agriculture and Food Research Council (U.K.)) is gratefully acknowledged.

REFERENCES

Addiscott, T.M. 1984. Modelling the interaction between solute leaching and intra-ped diffusion in clay soils. In Proceedings of the International Symposium on Water and Solute movement in Heavy Clay Soils, Wageningen, pp. 279-292.

Jarvis, N.J., and P.B. Leeds-Harrison. 1987a. Modelling water movement in drained clay soil I. Description of the model, sample output and sensitivity analysis. J. Soil Sci. 38: 487-498.

Jarvis, N.J., and P.B. Leeds-Harrison. 1987b. Modelling water movement in drained clay soil II. Application of the model in Evesham clay soil. J. Soil Sci. 38: 499-509.

Jarvis, N.J. 1989. 'CRACK' - A model of water and solute movement in cracking clay soils. Report 159, Division of Agricultural Hydrotechnics, Department of Soil Sciences, SLU, Uppsala, Sweden, 38pp.

Thomas, G.W., and R.E. Phillips. 1979. Consequences of water movement in macropores. J. Environ. Qual. 8: 149-152.

van Genuchten, M.T., and P.J. Wierenga. 1976. Mass transfer in sorbing porous media. I. Analytical solutions. Soil Sci. Soc. Am. J. 40: 473-480.

INTEGRATED ASSESSMENT OF EROSION AND NONPOINT-SOURCE POLLUTION IN A PALOUSE WATERSHED

Tony Prato and Merlyn Brusven[1]

ABSTRACT

Reduction in erosion and improvements in water quality at the outlet of a Palouse watershed were analyzed. Total net farm income and water quality increased when average erosion was reduced to 2T. When average erosion was reduced to 1T, water quality improved substantially but total net farm income declined.

The Palouse region of eastern Washington and northern Idaho is one of the most productive and highly eroding dryland wheat-producing regions of the world. The U.S. Soil Conservation Service has identified the Palouse region and associated streams as having severe erosion and water quality problems (USDA 1981). Cropland erosion in the Palouse exceeds 10 million metric tons per year (USDA 1984) and generates sediment and nutrients which pollute receiving waters and degrade fish spawning and rearing habitat. This paper examines the tradeoffs between improving water quality by reducing cropland erosion and net farm income in a Palouse watershed.

TARGET WATERSHED

The tradeoff analysis was done for the Tom Beall watershed which is located in the Lapwai Creek drainage of northern Idaho. This 4,563-hectare watershed contains 3,202 hectares of cropland. Due to steep and undulating topography, about 75% of the cropland in the watershed is classified as highly erodible. Annual average erosion is 27.8 tons per hectare per year (THY) based on current land use and farming practices. A winter wheat--spring pea rotation and conventional tillage with contour farming are the most common crop rotation and farming practice used in the watershed. High erosion rates cause considerable amounts of sediment, nutrients and pesticides to enter Tom Beall Creek, resulting in poor water quality (Brusven et al. 1986 and 1988).

EVALUATION PROCEDURES

The tradeoffs between reductions in erosion and total net farm income were determined using a linear programming optimization model. The model determined the optimal resource management systems (RMS's) for maximizing total net farm income for the 16 farms in the watershed subject to successively higher levels of erosion reduction. The AGNPS model (Young et al. 1985) was then used to determine the effects of optimal RMS's on water quality at the outlet of the watershed for storm events with four return periods: 10, 25, 50, and 100 years. Each storm was assumed to last 24 hours.

A Geographic Information System was used to assemble and analyze information on soil type, topography, watercourses, cropping pattern, watershed and field boundaries, conservation practices, and the movement of sediment and nutrients through the watershed. Soil erosion was calculated using the USLE (Wischmeier and Smith 1978). Eleven RMS's were evaluated: conventional tillage with up-and-down cultivation, cross-slope farming, contour farming, or divided

[1] Tony Prato, Professor of Agricultural Economics, and Merlyn Brusven, Professor of Entomology, College of Agriculture, University of Idaho, Moscow, ID.

slope farming; minimum tillage with cross-slope farming, contour farming, or divided slope farming; no tillage with cross-slope farming, contour farming, or divided slope farming; and permanent vegetation. A zero yield penalty was assumed for minimum tilled wheat and a 15% yield penalty for no-tilled wheat. Since peas are conventionally tilled, there was no yield penalty for peas.

RESULTS AND DISCUSSION

Cropland acreage shifted from conventional tillage to minimum tillage for erosion reductions up to 40%. No tillage replaced minimum tillage on highly eroding farms when total erosion was reduced 40-60%. To reduce erosion by more than 80% required extensive use of permanent vegetation. Minimum tillage with either cross-slope farming or contour farming was the most economically efficient RMS for reducing erosion.

Figure 1 illustrates the tradeoff between net farm income and erosion reduction. Reducing total erosion in the watershed by 40% of the baseline level caused total net farm income to increase 1.5% without cost sharing and 15.8% with cost sharing. The baseline level is the total erosion for a wheat-pea rotation using conventional tillage with contour farming. Total net farm income decreased 34.7% without cost sharing and 17.7% with cost sharing when total erosion was reduced 70%. Total net farm income declined rapidly beyond 40% erosion reduction without cost sharing and 60% erosion reduction with cost sharing.

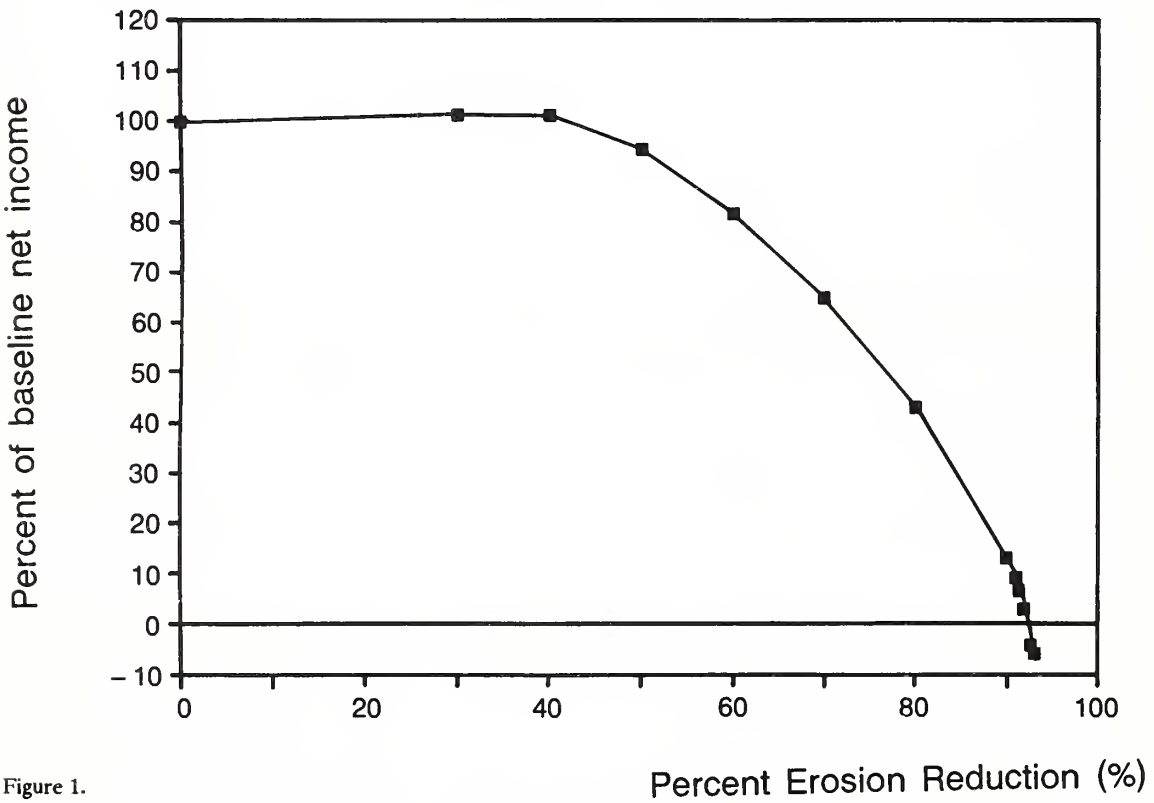


Figure 1.
Tradeoff between net farm income (without cost sharing) and erosion reduction for Tom Beall watershed.

Reducing average erosion in the watershed to 2T (11.2 THY) and 1T (2.3 THY) significantly reduced total and average sediment, nitrogen, and phosphorus at the outlet of Tom Beall Creek. Results are given in detail in table 1. Sediment, nitrogen and phosphorus levels increased with storm intensity, but at a decreasing rate. Between current practices and 2T, sediment declined 43-48%, and nitrogen and phosphorus dropped 34-38%. The lower limit in each percentage range corresponds to a 100-year storm and the upper limit corresponds to a 10-year storm. From 2T to 1T, sediment declined 48-53%, nitrogen declined 38-42%, and phosphorus decreased 40-45%. Average losses of sediment, nitrogen, and phosphorus for the four storm events decreased 45, 38, and 38%, respectively, with the optimal RMS's for 2T; and 72, 64, and 64%, respectively, with the optimal RMS's for 1T.

Table 1.
Total and average sediment, nitrogen, and phosphorus for alternative erosion control levels in Tom Beall watershed.

		Erosion Control Level					
		Current Practices		2T		1T	
	Event	Total[1]	Avg[2]	Total	Avg	Total	Avg
Sediment[3]	10	11,484	2.73	5,930	1.41	2,730	0.65
	25	17,979	4.28	9,884	2.35	4,801	1.14
	50	22,026	5.24	12,049	2.87	6,212	1.48
	100	26,074	6.20	14,066	3.56	7,624	1.81
Nitrogen[3,4]	10	19.11	4.55	12.00	2.86	6.92	1.65
	25	26.87	6.48	17.46	4.16	10.49	2.50
	50	31.16	7.41	20.57	4.89	12.56	2.99
	100	35.77	8.51	23.91	5.69	14.76	3.52
Phosphorus[3,4]	10	9.03	2.15	5.51	1.31	3.01	0.72
	25	12.80	2.72	8.05	1.92	4.66	1.11
	50	14.92	3.55	9.55	2.29	5.65	1.34
	100	17.09	4.07	11.16	2.65	6.68	1.59

[1]10^3 kilograms

[2]Tons per hectare for sediment and kilograms per hectare for nitrogen and phosphorus.

[3]Total equals amount released at watershed outlet. Average equals total divided by cropland acreage in watershed.

[4]Attached to sediment and dissolved in water.

CONCLUSION

Total net farm income and water quality in the Tom Beall watershed can be increased for moderate reductions in total erosion (down to 2T). Achieving the current soil erosion tolerance limit of 1T would greatly improve water quality at the watershed outlet but substantially reduce total net farm income.

LITERATURE CITED

Brusven, Merlyn A., Gene P. Carpenter. 1988. Nonpoint source pesticide contamination of waters in Idaho Palouse creeks and streams and its possible effects on aquatic organisms. Progress report submitted to Western Regional Pesticide Impact and Assessment Program.

Brusven, Merlyn A., Ron Rhew, and Roger Haro. 1986. Baseline monitoring analysis of Lapwai Creek and selected tributaries in relation to nonpoint source agricultural pollution. Research report submitted to Nez Perce Tribe, Lapwai, ID.

U.S. Department of Agriculture, Soil Conservation Service, Forest Service, Idaho Department of Health and Welfare, Division of Environment, Idaho Soil Conservation Commission, and Soil Conservation Districts. 1985. Detailed work outline, Idaho special studies, erosion control and water quality.

U.S. Department of Agriculture, Soil Conservation Service. 1984. Idaho's soil and water: Condition and trends, Boise, ID.

U.S. Department of Agriculture. 1981. Soil, water and related resources in the United States: Status, condition and trends, 1980 RCA appraisal, part I, Washington, DC.

Wischmeier, W.H., and D.D. Smith. 1978. Predicting rainfall erosion losses. Agricultural Handbook No. 537, U.S. Department of Agriculture, Washington, DC.

Young, Robert A., Charles A. Onstad, David D. Bosh, and Wayne P. Anderson. 1985. Agricultural nonpoint source pollution model, a guide to model users. North Central Soil Conservation Research Laboratory, Morris, MN.

PREDICTION OF CHEMICAL TRANSPORT IN AGRICULTURAL RUNOFF: AN INTEGRATED PROCESS APPROACH

Andrew N. Sharpley, S. J. Smith, L. R. Ahuja, and G. C. Heathman[1]

ABSTRACT

Soluble and particulate P, N, and pesticide transport in runoff is described by physically-based equations, detailed in this paper. Transport in solution is described by kinetic or equilibrium equations, with the parameters and constants a dynamic function of watershed management and the nature of the surface soil-rainfall interaction. Transport of sediment-bound chemicals is described by a relationship between enrichment ratio (chemical content of sediment/source soil) and soil loss. Predicted P and N concentrations of individual runoff events using these relationships compared well with values measured in runoff over a 10-year period from 9 grassed and 11 cropped watersheds in the Southern Plains (r^2 of 0.62 to 0.97) over a wide range in measured concentrations, soil types, and fertilizer application. The transport of alachlor, atrazine, and cyanazine in runoff from field plots in Iowa was also accurately described.

INTRODUCTION

Nonpoint-source pollution of lakes by agricultural runoff is one of the world's major water quality problems. Agricultural chemicals of primary concern are phosphorus (P), nitrogen (N), and pesticides. Long-term field studies can investigate the impact of agricultural management practices on the transport of these chemicals and corrective action can be taken when water quality criteria are exceeded. This can be inefficient and costly, however, with irreversible damage being done. Mathematical models have, thus, been developed describing the transport of agricultural chemicals in runoff (DeCoursey, 1985), with the purpose of aiding selection of management practices capable of reducing water quality problems. Although physically based descriptions of the various transport processes are used, a lack of needed information and limited field testing has resulted in oversimplifications.

This paper documents recent improvements made at our laboratory in descriptions of the rainfall-runoff interaction and nature of partitioning mechanisms between soluble, particulate (sediment-bound), and bioavailable forms of chemicals involved in the transport process. Model predictions are compared with field measurements of P and N transport in runoff from 20 unfertilized grassed and cropped watersheds in Oklahoma and Texas and pesticide transport from field plots in Iowa.

MODEL DESCRIPTION

Soluble Transport

Although a thin layer of surface soil (less than 10 mm) has been assumed to mix with rainfall during chemical transfer from soil to runoff. Ahuja et al. (1981) found that a chemical may be transferred to runoff from a soil depth as great as 20 mm. The degree of mixing and transfer of chemical from soil to rainwater, however, decreases exponentially with depth below the surface. An effective depth of complete and uniform mixing may be approximately assumed under high infiltration conditions (Ahuja et al. 1981, Ahuja 1982). This assumption was adequate for P,

[1] Soil Scientists; USDA-ARS, Water Quality and Watershed Research Laboratory, Durant, OK.

but not where low infiltration rates occur and for pesticides and non-adsorbed chemicals (Ahuja et al. 1981). Modeling of these two groups of chemicals is, thus, treated separately using uniform and non-uniform mixing models, respectively.

Kinetic - Uniform Mixing Model

The mean soluble P (SP) concentration of runoff ($\mu g\ L^{-1}$) during an event can be predicted by equation 1, which has been derived from an empirical relationship describing the kinetics of soil P desorption (Sharpley et al. 1981):

$$SP = \frac{K\ P_a\ E\ B\ t^{\alpha}\ W^{\beta}}{V} \qquad [1]$$

where P_a = initial available soil P content (mg kg^{-1})
E = effective depth of interaction between surface soil and runoff (mm)
B = bulk density of soil in this depth (kg m^{-3})
t = runoff event duration (min)
V = runoff volume (L)
W = water:soil ratio - approximated as runoff volume: suspended sediment content of runoff (L kg^{-1})
K, α and β = constants for a given soil.

For each runoff event, E was calculated from soil loss (kg ha^{-1}) using the following equation (Sharpley 1985a):

$$\ln(E) = i + 0.576\ \ln(\text{soil loss}) \qquad [2]$$

Non-Uniform Mixing Model

The model incorporates the varying degree of mixing between rain and soil water during chemical transfer to runoff, as well as the effects of infiltration on chemical movement into the soil before and after runoff initiation (Ahuja 1986). Numerical computations are made in small intervals of soil depth (1 mm) and time. Water is assumed to move into the soil in increments, equal to the amount needed to saturate a 1-mm depth. Starting from the time of runoff initiation, it is assumed that the degree of mixing, m, between rainfall and soil solution, decreases exponentially with soil depth.

$$m = \exp(-bz) \qquad [3]$$

where z = soil depth (mm)
b = constant

The concentration of chemical in runoff (Cr, mg L^{-1}) is given by:

$$C_r = \sum_{i=1}^{20} M_i\ C_i \qquad [4]$$

where i = soil interval
C_i = concentration of chemical in soil solution after mixing with an increment of rainfall (mg L^{-1}).

A more detailed description is given by Ahuja (1986).

Particulate Transport

The particulate P (PP), bioavailable P (BioP), and total N (TN) concentration of runoff were calculated from the enrichment of the respective nutrient forms in suspended sediment compared to surface soil:

$$\text{Runoff PP} = (\text{Soil TP})(\text{Sediment concentration})(\text{ER}) \quad [5]$$

$$\text{Runoff BioP} = (\text{Soil BioP})(\text{Sediment concentration})(\text{ER}) \quad [6]$$

$$\text{Runoff TN} = (\text{Soil TN})(\text{Sediment concentration})(\text{ER}) \quad [7]$$

where soil TP = soil TP content (mg kg^{-1})
soil BioP = soil bioavailable P content (mg kg^{-1})
soil TN = soil TN content (mg kg^{-1})
ER = enrichment ratio

Enrichment ratio is predicted from soil loss (kg ha^{-1}) using the following relationship developed by Sharpley (1985b):

$$\ln(\text{ER}) = 1.21 - 0.16 \ln(\text{soil loss}) \quad [8]$$

EXPERIMENTAL

Study Area

Management of 20 watersheds representing major agricultural practices of the Southern Plains area of Oklahoma and Texas has been detailed (Sharpley et al. 1985). Suffice it to note here that they encompassed a range of sizes (1-122 ha), soils (Inceptisols, Mollisols, and Vertisols), slopes (1-9%), grasses, crops (cotton, oats, sorghum, and wheat), fertilizer application rates (N, 0-95 and P, 0-23 kg ha^{-1} yr^{-1}), and study periods (6-10 yr).

Watershed runoff was measured using precalibrated flumes or weirs, with flow-weighted samples collected from each runoff event (Sharpley et al. 1982). Runoff samples were refrigerated at 4 C until analysis. Surface soil samples (0-50 mm depth) were collected at four sites in each watershed (near the flumes) at monthly intervals and composited. The samples were then air-dried and sieved (2 mm).

Nutrient concentrations of runoff and soil were determined by analytical procedures noted previously (Sharpley et al. 1985).

RESULTS

Soluble P

The mean SP concentration of each runoff event was calculated using equation 1. Values of equation 3 constants were calculated from the ratio of percent clay to organic carbon content of surface soil at each watershed location (Sharpley 1983). Event duration was set at 30 min, an approximate value for a representative runoff event. Soil bulk densities were obtained from field measurements (1.40, 1.35, and 1.40 Mg m^{-3} for Houston Black, Kirkland, and Woodward soils, respectively), and soil available P content measured before each runoff event. For all watersheds, measured and predicted SP values were statistically similar (1.0% level), as determined by analysis of variance (table 1). Coefficients of determination and slopes of the regression between

measured and predicted values at each watershed were close to 1.00, with intercept values close to zero.

Particulate P and N

The PP, BioP, and TN concentrations in runoff from the watersheds were predicted using equations 5-8, soil loss and concentration of each runoff event, and TP, BioP, and TN content of the surface soil before runoff.

Measured and predicted PP and TN values were statistically similar (at 1% level). Similarly, BioP concentrations of runoff from watersheds at El Reno and Woodward during 1985 and 1986 were closely predicted, covering a wide range of measured values (4-3596 $\mu g\ L_{-1}$) (fig. 1).

Predictions of particulate nutrient transport were less accurate for grassed compared to wheat watersheds (table 1 and fig. 1). This may result from lower soil loss from grass and use of the same values of equation 8 constants for all watersheds.

Table 1.
Measured and predicted soluble P, particulate P, and total N concentrations in runoff events averaged for the study period.

Water-shed	Number of events	Soluble P			Particulate P			Total N		
		Meas.	Pred.	r^2	Meas.	Pred.	r^2	Meas.	Pred.	r^2
		$\mu g\ L^{-1}$			$\mu g\ L^{-1}$			$\mu g\ L^{-1}$		
Grass										
FR1	32	122	133	0.85	96	64	0.67	4.98	1.89	0.63
FR2	38	167	168	0.98	121	119	0.95	2.61	1.76	0.60
FR3	37	100	104	0.98	103	77	0.82	1.81	0.85	0.52
FR4	33	168	169	0.98	90	90	0.64	2.41	1.82	0.57
Y14	31	104	105	0.97	289	303	0.93	1.83	1.39	0.86
W10	21	106	108	0.94	95	76	0.73	1.54	1.06	0.62
SW11	27	215	139	0.74	685	545	0.81	2.68	2.23	0.82
W1	27	173	168	0.95	565	476	0.65	3.07	2.20	0.68
W2	59	201	194	0.96	1329	1324	0.82	5.32	4.35	0.66
Wheat										
FR5	68	282	268	0.97	2247	2077	0.86	10.70	9.86	0.94
FR6	51	345	338	1.00	1017	1759	0.91	10.90	10.26	0.97
FR7	47	589	596	0.97	795	634	0.61	5.98	4.39	0.63
FR8	55	424	415	1.00	1224	1191	0.96	7.71	6.96	0.91
W3	41	463	462	0.99	11134	10487	0.99	43.37	41.60	0.99
W4	66	925	950	0.97	1400	1284	0.80	6.68	5.76	0.82
Mixed Crops and Grass										
Y	54	173	176	0.94	396	388	0.85	2.32	1.99	0.63
Y2	38	137	142	0.89	534	536	0.93	2.31	2.33	0.78
Y6	24	58	63	0.73	1631	1547	0.86	5.69	5.27	0.93
Y8	21	6	34	0.83	927	719	0.69	3.07	2.69	0.95
Y10	31	57	56	0.71	1815	1740	0.87	6.05	6.29	0.93

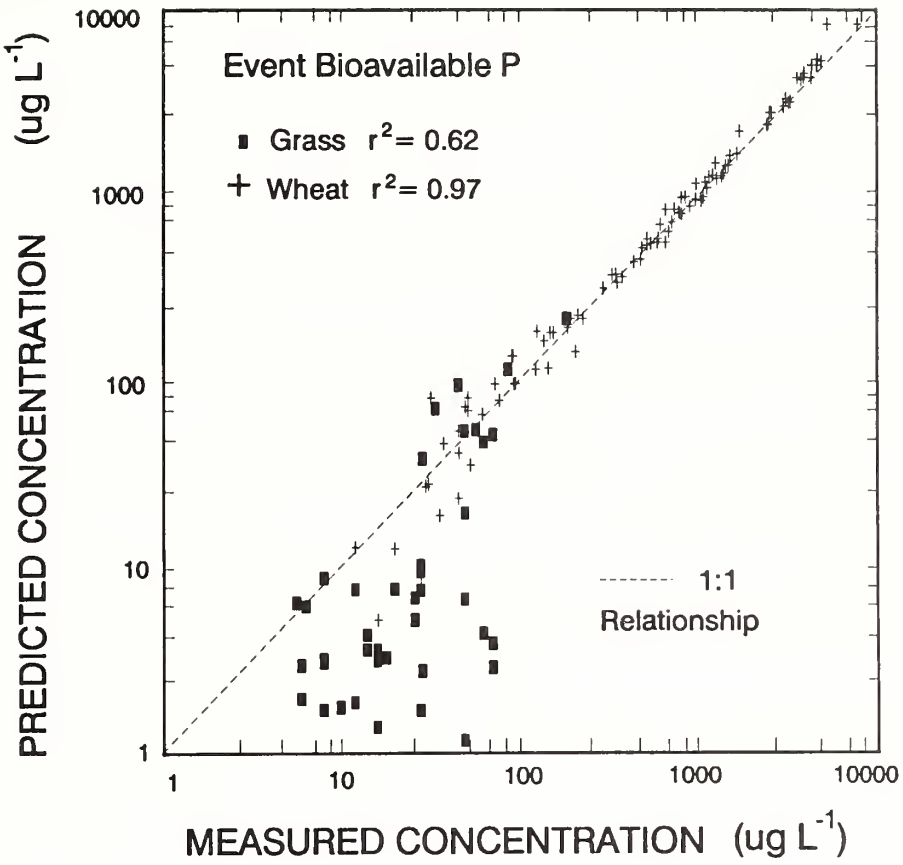


Figure 1.
Relationship between measured and predicted bioavailable P concentration in individual runoff events.

Pesticide Transport

The non-uniform mixing model was applied to the field-plot data of Baker et al. (1982) for pesticide transport in runoff (fig. 2). The model gave a good fit to the data for runoff from bare plots of Clarion sandy loam. Accurate predictions were also obtained for transport in runoff from plots with surface residue (750 kg ha^{-1}) (Heathman et al. 1986). The value of the desorption constant for a given chemical, was generally constant among replicates of a given residue application. The constant describing degree of rain - soil water mixing, varied between replicates due to spatial variability of surface conditions. In general, residue cover increased the best-fit values of the two parameters, indicating that residue cover shielded some of the chemical on the soil surface from mixing with rainfall and decreased the degree of mixing with depth, by reducing raindrop kinetic energy.

DISCUSSION

It is clear from the results of the present study that the kinetic or uniform mixing, enrichment ratio, and non-uniform mixing models gave accurate predictions of soluble and particulate chemicals transported in runoff from a range of agricultural management practices. The data indicate, however, that further model development is warranted. For the kinetic uniform mixing

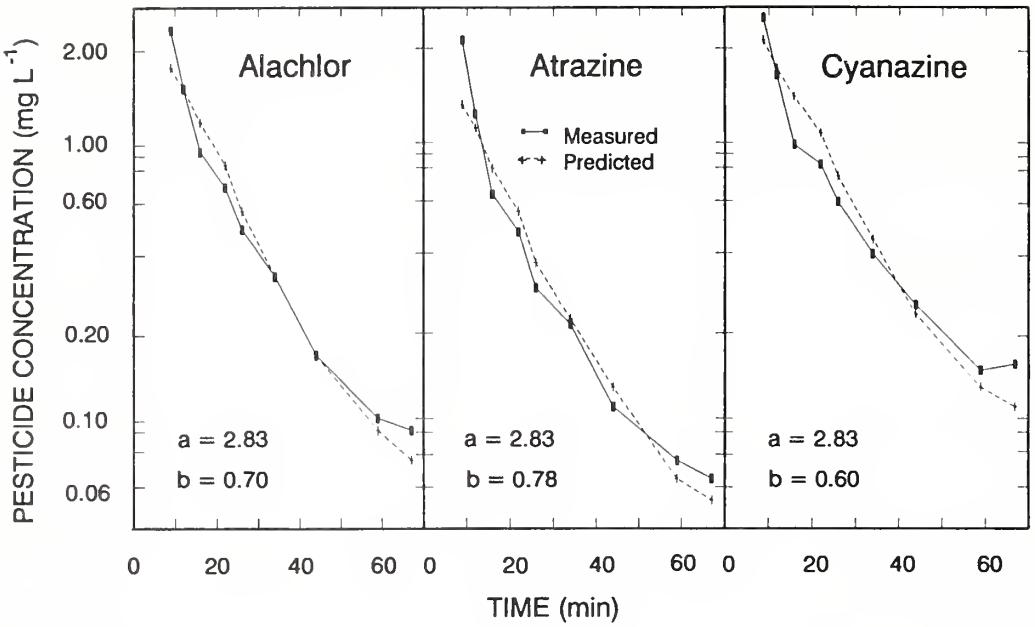


Figure 2.
Relationship between measured and predicted alachlor, atrazine and cyanazine concentration in runoff from field plots in Iowa (adapted from Heathman et al. 1986).

model, this will involve making depth of rainfall-surface soil interaction (E, equation 1), a function of agricultural management practice such as soil tillage, which will allow a wider application of the model. This will be particularly important for reduced or no-till systems, where crop residue will influence the degree of mixing between rainfall and surface soil. Additionally, chemical release from the crop residue will need to be considered.

It appears that the influence of crop residue on the transfer of a chemical from soil to runoff, may be partially accounted for in the non-uniform mixing model. Further development should account for interflow or subsurface runoff which may be significant in sloping soils with surface horizons of high permeability underlain by a horizon of much lower permeability and in soils with well-formed root channels and other macropores. This water has a high potential for transporting soil chemicals in solution in runoff.

Further development of the predictive relationships for particulate chemical transfer (equations 5-8) will involve making the slope and intercept values of equation 8, a function of factors affecting soil loss or runoff energy. These factors include rainfall intensity and duration, crop residue cover, and management practices. By accounting for these factors, the prediction of and, thus, particulate P, BioP, and total N transport in runoff should be improved, particularly for low flow events.

REFERENCES

Ahuja, L.R. 1982. Release of soluble chemical from soil to runoff. Trans. ASAE 25:948-956, 960.

Ahuja, L.R. 1986. Characterization and modeling of chemical transfer to runoff. Adv. Soil Sci. 4:149-188.

Ahuja, L.R., A. N. Sharpley, M. Yamamoto, and R.G. Menzel. 1981. The depth of rainfall-runoff-soil interaction as determined by ^{32}p. Water Resour. Res. 17:969-974.

Baker, J.L., J.M. Laflen, and R.O. Hartwig. 1982. Effects of corn residue and herbicide placement on herbicide runoff losses. Trans. ASAE 25:340-343.

DeCoursey, D.G. 1985. Mathematical models for nonpoint source pollution control. J. Soil Water Conserv. 40:408-413.

Heathman, G.C., L.R. Ahuja, and J.L. Baker. 1986. Test of a non-uniform mixing model for transfer of herbicides to surface runoff. Trans. ASAE 29:450-455, 461.

Sharpley, A.N. 1983. Effect of soil properties on the kinetics of phosphorus desorption. Soil Sci. Soc. Am. J. 47:462-467.

Sharpley, A.N. 1985a. Depth of surface soil-runoff interaction as affected by rainfall, soil slope, and management. Soil Sci. Soc. Am. J 49:1010-1015.

Sharpley, A.N. 1985b. The selective erosion of plant nutrients in runoff. Soil Sci. Soc. Am. J. 49:1527-1534.

Sharpley, A.N., S.J. Smith, and R.G. Menzel. 1982. Prediction of phosphorus losses in runoff from Southern Plains watersheds. J. Environ. Qual. 11:247-251.

Sharpley, A.N., L.R. Ahuja, M. Yamamoto, and R.G. Menzel. 1981. The kinetics of phosphorus desorption from soil. Soil Sci. Soc. Am. J. 45:493-496.

Sharpley, A.N., S.J. Smith, W.A. Berg, and J.R. Williams. 1985. Nutrient runoff losses as predicted by monthly and annual soil sampling. J. Environ. Qual. 14:354-360.